CONTOUR, COLOR AND SHAPE ANALYSIS
BEYOND THE STRIATE CORTEX

ROBERT DESIMONE¹, STANLEY J. SCHEIN², JEFFREY MORAN¹ and LESLIE G. UNGERLEIDER¹
¹Laboratory of Neuropsychology, National Institute of Mental Health, Section on Visual Processing,
National Eye Institute, National Institutes of Health, Bethesda, MD 20805, U.S.A.

Abstract—The corticocortical pathway from striate cortex into the temporal lobe plays a crucial role in
the visual recognition of objects. Anatomical studies indicate that this pathway is mainly organized as
a serial hierarchy of multiple visual areas, including V1, V2, V3, V4, and inferior temporal cortex (IT).
As expected from the anatomy, we have found that neurons in V4 and IT, like those in V1 and V2, are
sensitive to many kinds of information relevant to object recognition. In the spatial domain, many V4
cells exhibit length, width, orientation, direction of motion and spatial frequency selectivity. In the spectral
domain, many V4 cells are also tuned to wavelength. Thus, V4 is not specialized to analyze one particular
attribute of a visual stimulus; rather, V4 appears to process both spatial and spectral information in
parallel. A special contribution of V4 neurons to visual processing may lie in specific spatial and spectral
interactions between their small excitatory receptive fields and large silent suppressive surrounds. Thus,
although the excitatory receptive fields of V4 neurons are small, the responses of V4 neurons are influenced
by stimuli throughout a much larger portion of the visual field. In IT, neurons also appear to process
both spatial and spectral information throughout a large portion of the visual field. However, unlike V4
neurons, the excitatory receptive fields of IT neurons are very large. Many IT neurons, for example, are
selective for the overall shape, color, or texture of a stimulus, anywhere within the central visual field.
Together, these results suggest that within the areas of the occipito-temporal pathway, many different
stimulus qualities are processed in parallel, but the type of analysis may become more global at each stage
of processing.

INTRODUCTION

Over a dozen separate visual areas have been described in the extrastriate cortex of the macaque,
including areas V2, V3, VP (or ventral V3), V3A, V4, MT, MST, VIP, POa, PO, STP, IT and ill-defined
areas within TF, TH, PG and the frontal lobe. What role does each of these areas play in the recognition
of visual stimuli? One prevalent view is that the primary visual cortex sends visual information di-
rectly or indirectly to many visual areas and these areas process different attributes of a stimulus in
parallel with each other. We suggest an alternative view, namely, that only a few of the multiple visual
areas play a crucial role in the recognition of objects, and these areas are primarily organized in a serial
hierarchy. Within each of the areas several different stimulus qualities are processed in parallel, but the
type of analysis may become more global at each stage of processing. In this paper we will summarize
the neurobehavioral evidence for a cortical object-recognition system, outline the anatomical or-
ganization of this system, and describe the results of recent electrophysiological studies of two of the
major areas in this system, areas V4 and IT.

THE CORTICAL SYSTEM FOR OBJECT
RECOGNITION

As Ungerleider and Mishkin (1982) have recently
reviewed the evidence for a cortical object-
recognition system, we will only briefly summarize
the evidence here. According to the Ungerleider and
Mishkin model, striate cortex is the source of two
corticocortical pathways, each involving several vi-
sual areas. One of the pathways is directed ventrally
into the temporal lobe and is crucial for object
recognition. The other is directed dorsally into the
parietal lobe and is crucial for spatial perception
and visuomotor performance. The primary evidence
for this distinction comes from the contrasting effects
of temporal and parietal lesions in monkeys (Pohl, 1973;
Iwai and Mishkin, 1968; Brody and Pribram, 1978;
Mishkin and Ungerleider, 1982; also see for reviews,
Mishkin, 1972; Gross, 1973; Dean, 1982; Ungerleider
and Mishkin, 1982; Mishkin et al., 1983). Lesions of
inferior temporal cortex (IT) cause severe deficits in
performance of a wide variety of visual discrimi-
nation but not visuo-spatial tasks. Monkeys with IT
lesions are impaired, for example, in learning pattern
discrimination performance, and object discriminations and in distinguishing be-
etween a familiar and a novel object. By contrast,
posterior parietal lesions do not affect visual discrimi-
nation performance, but instead cause severe deficits
in visuo-spatial performance, such as visually-guided
reaching and judging which of two identical objects
is located closest to a visual landmark. Physiological
evidence also supports this distinction, as neurons in
IT (Desimone et al., 1984) but not parietal cortex

441
(Robinson et al., 1978) are highly sensitive to the shape or color of a stimulus while neurons in parietal cortex are more sensitive to direction of stimulus motion and tracking eye movements (Lynch et al., 1977; Robinson et al., 1978; Mountcastle et al., 1981; Newsome and Wurtz, 1982; Sakata et al., 1983). Although pattern vision and visuo-spatial functions are commonly associated with the geniculostriate and tectofugal systems, respectively, in the monkey it has been shown that both pattern recognition and visuo-spatial functions are dependent on striate cortex. Lesions of striate cortex (but not the superior colliculus) severely impair both pattern and spatial vision (Mishkin and Ungerleider, 1982) and nearly eliminate visually driven activity in both IT (Rocha-Miranda et al., 1975) and posterior parietal cortex (M. E. Goldberg, personal communication). Thus, IT appears to be part of a system that originates in striate cortex and is necessary for recognizing objects while posterior parietal cortex is part of a system that also originates in striate cortex but is necessary for appreciating the spatial relationships among objects.

ANATOMY OF THE OBJECT-RECOGNITION SYSTEM

Although early anatomical experiments indicated that the corticocortical pathways from striate cortex into the temporal and parietal lobes must involve visual areas located in prestriate cortex (Kuypers et al., 1965; Jones and Powell, 1970), it has only been recently that many of the prestriate areas and their connections have been delineated. Surprisingly, the occipito-temporal system apparently involves relatively few visual areas while the occipito-parietal system involves many. Both systems begin with the outputs of striate cortex, which are predominantly to areas V2, V3 and MT (Cragg and Ainsworth, 1969; Zeki, 1969; Ungerleider and Mishkin, 1979; Weller and Kaas, 1983; Ungerleider, 1985). Since V2 and V3 are associated with both systems, while MT is associated primarily with the occipito-parietal system, the projections to V2 and V3 will be described first.

The largest projection field of striate cortex is V2, a visuostructically organized area that nearly surrounds V1. There is, in addition, a much smaller and weaker projection from the lower visual field representation of V1 to the lower field representation of V3, a narrow area bordering V2. (For a variety of reasons, some consider the upper field representation in V3 to be a separate area from the lower field representation and term the former “VP” (Newsome et al., 1980; Felleman et al., 1984). While we feel this question is still open, we lean towards considering V3 to be a single visual area that may have somewhat different connections and neural properties in the upper and lower portions of the visual field. Similarly, though striate cortex may also have somewhat different connections (Zeki, 1980b) and properties (Poggio et al., 1975; Zeki, 1983a; Livingston and Hubel, 1984) in different portions of the visual field, it is regarded as a single area.)

The projection from V1 to V2, V3 and MT originates primarily from V1 cells located above layer 4C and terminates predominantly in layer 4 of each of these areas (Tigges et al., 1973, 1974; Wong-Riley, 1978; Rockland and Pandya, 1979; Ungerleider and Mishkin, 1979; Lund et al., 1981; Maunsell and Van Essen, 1983c; Weller and Kaas, 1983; Ungerleider, 1985). The reciprocal projection from each of these areas back to V1 originates primarily from infragranular cells and terminates in layers excluding layer 4C (Tigges et al., 1973, 1974; Wong-Riley, 1978; Rockland and Pandya, 1979; Lund et al., 1981; Ungerleider et al., 1982; Maunsell and Van Essen, 1983c). This general pattern of “forward” projections from the supragranular layers to granular layer 4 characterizes not only the direct projections of V1, but projections of extrastriate areas similarly directed away from V1 (Tigges et al., 1974; Rockland and Pandya, 1979; Lund et al., 1981; Ungerleider et al., 1982; Maunsell and Van Essen, 1983c). Likewise, the “backward” pattern of projections from the infragranular layers to layers excluding granular layer 4 characterizes projections directed towards V1 (Tigges et al., 1974; Rockland and Pandya, 1979; Ungerleider et al., 1982, 1983; Maunsell and Van Essen, 1983c). Recently, Maunsell and Van Essen (1983c) have also described an “intermediate” laminar pattern of connections that characterizes projections directed neither towards nor away from V1. Thus, the laminar pattern of connections between areas helps to establish the direction of the flow of information through the cortex and thereby the hierarchical arrangement of cortical visual areas (Rockland and Pandya, 1979; Maunsell and Van Essen, 1983c; Van Essen and Maunsell, 1983). A similar principle may be followed in other sensory systems (see Friedman, 1983).

The major forward projections of areas V2 and V3 are to V4 and MT (Zeki, 1971; Maunsell and Van Essen, 1983c; Ungerleider et al., 1983; Ungerleider, 1985). Dorsally, V4 borders area V3 and receives projections from the lower field representations of V2 and V3, as originally described by Zeki (1971). Although Zeki did not investigate the cortex on the ventral surface of the hemisphere, we consider the cortex adjacent to V3 on the ventral surface to be the upper field representation of V4, as it receives projections from the upper field representations of V2 and V3 (Ungerleider et al., 1983; Sousa et al., 1984). Dorsally and ventrally, the visuostructural organization of V4 roughly parallels that of V2 and V3, in that the central representation of V4 is located laterally in the hemisphere and the peripheral representation is represented medially (Gattass et al., 1983; Ungerleider, 1985).

Although there is at least a crude visuostructural organization within V4, at least with regards to upper vs lower and central versus peripheral visual fields, the details of this organization appear to be complex.
On the one hand, Van Essen and Zeki (1978) describe multiple representations of the vertical meridian within V4, and Maguire and Baizer (1984) parcel the dorsal portion of V4 into at least two partial representations of the lower visual field. On the other hand, Gattass et al. (1985) report that V4 contains a single, but crude, visual field representation, and Ungerleider et al. (1983) report that single locations within V2 generally project to single locations within V4, suggesting that V4 could be a single area. Until more is known, we will consider "V4" to be a single area with a complex organization.

While the central representation of V4 has minor projections to both MT and a small zone in the parietal cortex, its major output is to IT (Desimone et al., 1980; Felleman and Van Essen, 1983; Ffennermacher et al., 1984; and unpublished data). Both the lower and upper visual field representations in V4 project to IT. Physiological studies have shown that, unlike the visual areas that precede it, IT has no discernible visuotopic organization (Desimone and Gross, 1979). Rather, IT neurons have very large receptive fields that nearly always include the center of gaze and frequently cross the vertical midline into the ipsilateral visual field. IT appears to be the "last" visual area in the cortical system for object recognition, as its cortical outputs are to areas in the temporal and frontal lobe that are probably not exclusively concerned with vision (Kuypers et al., 1965; Jones and Powell, 1970).

In summary, the pathway from striate cortex into the temporal lobe begins with projections to areas V2 and V3. These areas project in turn to area V4, and V4 projects into the inferior temporal cortex. This anatomical sequence is illustrated diagramatically in Fig. 1. As there is a projection from striate cortex to both V2 and the dorsal portion of V3, and projections from both V2 and V3 to V4, there is a possibility of parallel processing within this system. However, the dominant organization of the occipito-temporal pathway appears to be a serial one. The remainder of the extrastriate visual areas seems to be more closely associated with the posterior parietal than the inferior temporal cortex, and we will review their connections here only briefly.

The major forward connections of MT are to areas in the superior temporal and intraparietal sulci that project in turn to cortex of the inferior parietal lobule (Ungerleider et al., 1982; Maunsell and Van Essen, 1983c). MT also has an "intermediate" type laminar connection with area V4, but this connection is weak and inconsistent (Ungerleider et al., 1982; Ffennemn and Van Essen, 1983; Maunsell and Van Essen, 1983c). Thus, MT receives its inputs from areas that participate in the occipito-temporal pathway but its outputs so far appear to be mainly directed into the
parietal lobe. Consistent with this conclusion, MT lesions cause impairments in tracking eye-movements and visually-guided hand movements but have no effect on pattern discrimination (Newsome et al., 1983; Gross, Gattass, and Barber, personal communication; Ungerleider and Mishkin, unpublished data).

There are additional visual and polysensory areas in extrastriate cortex that appear to lie outside the occipito-temporal pathway. Areas V3A and PO in dorsal prestriate cortex are connected with the peripheral but not the central visual field representations of several other areas, and therefore would appear to play a minor role in object recognition (Zeki, 1980b; Colby et al., 1983; Ungerleider, unpublished data). Other areas such as STP in the superior temporal sulcuses, parts of TF and TH on the parahippocampal gyrus, and the dorsolateral prefrontal cortex receive polysensory inputs and have strong interconnections with the parietal but not the inferior temporal cortex (for review. see Gross et al., 1981). Finally, the frontal eye fields receive inputs from multiple prestriate areas and are concerned with eye movements rather than object recognition (Goldberg and Bushnell, 1981).

NEURAL PROPERTIES IN THE VISUAL AREAS OF THE OCCIPITO-TEMPORAL PATHWAY

Given the nearly sequential anatomical route to inferior temporal cortex from striate cortex, one would expect many types of visual information relevant to pattern recognition to be processed within each area along the route. Indeed, electrophysiological studies have shown that neurons in striate cortex, V2 and V3 are sensitive to one or (usually) more stimulus qualities such as length, width, orientation, direction of motion, spatial frequency, disparity, texture, and color (Dow, 1974; Poggio et al., 1975; Schiller et al., 1976a, 1976b; Baizer et al., 1977; Baizer, 1982; De Valois et al., 1982a, 1982b). In striate cortex and possibly V2, many of these stimulus qualities are represented within separate columnar or laminar systems (Dow, 1974; Michael, 1981; Tootell et al., 1983; Livingston and Hubel, 1984). Thus, to the extent that there is any segregation of stimulus qualities such as color or orientation in the cortex, the segregation may take place within visual areas rather than between areas. The apparent exceptions to this scheme of multidimensional processing within each visual area have been areas MT and V4. MT cells are sensitive to direction of stimulus motion but not color (Zeki, 1974, 1978, 1980a, 1983a). However, MT is associated with the occipito-parial system, and we expect the areas in this system to process different kinds of visual information from the visual areas in the occipito-temporal system. Moreover, even within the occipito-parial system processing may be multidimensional. Neurons in MT, for example, are sensitive to velocity, disparity and orientation, in addition to direction of motion (Maunsell and Van Essen, 1983a, 1983b; Albright, 1984). The other apparent exception to multidimensional processing, area V4, was originally reported to be specialized almost exclusively for the analysis of color (Zeki, 1973, 1974, 1978, 1980a). This claim was surprising, since V4 provides the exclusive input to IT, and IT neurons are sensitive to many stimulus qualities in addition to color. Later, several groups reported that most V4 cells were not highly selective for color (Kruger and Gouras, 1980; Fischer et al., 1981; de Monasterio and Schein, 1982; Schein et al., 1982), which left the status of V4 in question. In the remainder of this review, we will first describe some of our recent experiments on the spatial and spectral properties of V4 cells, experiments which indicate that V4 also processes information in parallel. Then we will summarize the properties of neurons in IT and suggest how they may contribute to the global analysis of shape and, ultimately, object recognition.

SPATIAL PROPERTIES OF V4 RECEPTIVE FIELDS

We recorded from over 200 single neurons within the representation of the central 5' in V4 (see Fig. 1), in anesthetized (N2O), paralyzed macaques (Desimone and Schein, 1983). Because neurons in V4 often have a variable response, all cells were studied with randomized presentation of stimuli and responses were averaged by computer.

We found receptive fields in V4 to be several times larger than in striate cortex. At an eccentricity of 1', V4 receptive fields are about 4 times larger than striate receptive fields in linear dimension, or 16 times larger in area, and at an eccentricity of 3' they are about 6 times larger in linear dimension, or 36 times larger in area. Thus, the properties of cells in V4 must ultimately derive from those of many cells in striate cortex.

Although receptive fields in V4 are larger than those in striate cortex, this does not mean that V4 cells are necessarily less spatially selective. Using quantitative measures of selectivity, we found that the proportions of cells in V4 sensitive to the length, width and direction of motion of a bar are comparable to the proportions of cells reported by others (Schiller et al., 1976a; De Valois et al., 1982a) in striate cortex. The distribution of orientation tuning bandwidths is also overlapping in the two areas (as narrow as 25' at half-maximum response in V4), although there appears to be a somewhat greater proportion of unoriented cells in V4. Likewise, the distribution of peak spatial frequencies in V4, ranging from 0.12 to 8 c/deg, almost completely overlaps that reported in striate cortex (De Valois et al., 1982b), although there appear to be more cells in V4 tuned to low spatial frequencies. Thus, while there are some quantitative differences between the spatial properties of cells in V4 and striate cortex, there appears to be
Fig. 2. Receptive fields of three representative V4 neurons. At the top are diagrams of the light and dark excitatory zones. At the bottom are the response curves. Stimuli were narrow light and dark bars, flashed at multiple positions, with separations indicated by the scale bars below each diagram. Responses were calculated from the average firing rate for five stimulus presentations, less the average spontaneous firing rate. For all three cells the dark and light excitatory zones were overlapping, and in this respect they resemble complex cells in striate cortex. For some cells in V4, the peaks of the dark and light zones were spatially offset (cell 1), but for most cells the zones were coextensive (cells 2 and 3). Some cells had a preference for either light or dark (cell 2), but for other cells the responses were nearly equivalent (cell 3).

a similar degree of spatial selectivity in both areas. Comparable spatial selectivity has also been reported for area V2 in the macaque (Baizer et al., 1977; Foster et al., 1983).

The structure of excitatory receptive fields in V4 can also be compared with the structure of excitatory receptive fields in striate cortex. Examples of receptive fields that were mapped quantitatively with narrow light and dark flashed bars are shown in Fig. 2. We found that almost all receptive fields in V4 have overlapping dark and light excitatory zones. The dark and light excitatory zones are not always equivalent: some cells have a preference for either light or dark, and in some cells the peaks of the dark and light zones may even be displaced. Yet, the zones are almost always at least partially overlapping, so in this respect V4 receptive fields resemble those of complex cells in striate cortex.

In other respects, the responses of cells in V4 can span the full range of characteristics reported for both simple and complex cells in striate cortex. For example, when we study the responses of V4 cells to moving bars of different widths, the response of many cells increases as bar width is increased up to the width of the excitatory receptive field, while the response of other cells peaks for a bar much narrower than the excitatory receptive field. Thus, as in striate cortex (Hubel and Wiesel, 1962, 1968; Schiller et al., 1976a), the evidence suggests that in some V4 cells there is primarily summation within the receptive field, while in other cells there may be primarily antagonistic mechanisms.

Likewise, when we tested V4 cells with drifting gratings we found a similar range of properties. At one extreme are cells that respond best to a low frequency grating, with a half-cycle that is equal to or wider than the receptive field, suggesting summation within the field. At another extreme are cells that respond best to a high frequency grating, with many cycles within the receptive field, suggesting antagonistic mechanisms within the field. Examples of both types are shown in Fig. 3. The cells that respond best to low frequencies tend to give more modulated responses to the optimal drifting grating than do cells that respond best to high frequencies. In striate cortex, summation and modulation are characteristics of simple cells, while antagonism within a uniform response zone and failure to modulate are characteristics of complex cells (Maffei and Fiorentini, 1973; Schiller et al., 1976b; Movshon et al., 1978a, 1978b; De Valois et al., 1982b). In V4, these properties appear to lie along a continuum.

A complete range of phase sensitivities is also seen in V4. When V4 cells are tested with flashed gratings of optimal spatial frequency but varying phase, some cells respond to only a narrow range of phases, and thus appear to respond linearly. Some cells have a small response at an opposite phase as well, while others have two best phases that are 180° opposite. Finally, some cells show virtually no phase sensitivity at all. In striate cortex, it has been reported that sensitivity to a single phase is a characteristic of a simple cell, while absence of phase sensitivity is a characteristic of a complex cell (Maffei and Fiorentini, 1973; Movshon et al., 1978a, 1978b; De Valois et al., 1982b). If so, then the continuum of phase sensitivity in V4 also appears to span the range of simple and complex behavior. [Alternatively, phase sensitivity may be represented along a continuum among cells in striate cortex as well (Dean and Tolhurst, 1983; Hochstein and Spitzer, 1983).]
suggested an even closer similarity between the two areas.

While we are not able to classify cells as purely simple or purely complex in V4, many V4 cells could be placed along a continuum from simple-like to complex-like based on the degree of summation, modulation, or phase sensitivity they exhibit. What is the significance of differences along these particular response dimensions? We speculate that high-frequency, phase insensitive cells with relatively large receptive fields might be particularly useful for the analysis of texture, while low frequency cells with greater phase sensitivity might be more useful for the analysis of borders and contours. A similar suggestion has been made for cells in striate cortex (Hammond and Mackay, 1977), and these results suggest that such a separation of function may be carried forward in V4.

To summarize, although the receptive fields of cells in V4 are many times larger than in V1, there is a high degree of spatial selectivity in both areas. Not only do V4 cells appear to continue the analysis of length, width, orientation and spatial frequency begun in striate cortex, but they also span the range of simple cell to complex cell behavior. Not surprisingly, V4 cells also appear to continue the analysis of spectral information.

**SPECTRAL PROPERTIES OF V4 RECEPTIVE FIELDS**

When we measured the response of V4 cells to white stimuli and to blue, green, yellow, and red broad-band colored stimuli, we found that most cells gave at least a small response to all stimuli (Schein et al., 1983). Across cells, the average response to the least effective color stimulus tested was 20% of the response to the best stimulus, and the average response to white light was 60% of the response to the best color stimulus. Thus, we have found, as have other studies, that V4 cells are not narrowly tuned color "detectors" (Van Essen and Zeki, 1978; Kruger and Gouras, 1980; Fischer et al., 1981; de Monasterio and Schein, 1982; Schein et al., 1982).

Nonetheless, the vast majority of V4 cells show evidence of spectral selectivity. When tested with narrow-band stimuli equated for either photopic luminosity or number of quanta, most V4 cells have a response that peaks at some wavelength and falls to spontaneous activity at others. The spectral bandwidths of the majority of the cells are narrower than the cones and are comparable to those of the color-opponent cells in the retina and lateral geniculate body (de Monasterio and Schein, 1982). The results from the cell illustrated in Fig. 4 are representative of most V4 cells. The stimuli were narrow-band spectral lights, ranging in wavelength from 460 nm (blue) to 660 nm (red). The cell gave an excitatory response throughout most of the spectrum but was wavelength sensitive, with a peak response to the 500 nm stimulus, in the blue-green part of the spectrum. The cell did not respond to wavelengths from 580 to 660 nm, in the yellow-red part of the spectrum, but it did give a good response to white light, not unlike some of the broadly-tuned color-opponent cells in the lateral geniculate body (Creutzfeldt et al., 1979). There were no explicitly color-opponent responses, such as inhibition or off-responses to stimuli in the red part of the spectrum. Although some cells in V4 do give color-opponent responses or fail to respond to white light, these cells are not common.

A broad spectral response curve, an absence of color-opponent responses and a good response to white light seem in many ways analogous to proper-
ties of a broad-band color filter, such as a piece of colored glass. Such a filter simply lets energy at some wavelengths pass more efficiently than at others and always lets some energy through from a white light, as white contains all wavelengths. If V4 cells acted like a photocell covered with a broad-band filter, we would expect them to respond not only to white light but also to other broad-band lights that vary in primary color but contain energy at many wavelengths. Depending on the criteria used, such a “color-filter” cell may or may not be classified as “color-selective.” Thus, our results may help explain the wide differences in reported incidence of color-selective cells in V4 (Zeki, 1973, 1978, 1980a, 1983a; Van Essen and Zeki, 1978; Kruger and Gouras, 1980; Fischer et al., 1981; Schein et al., 1982). Whether or not V4 cells are considered to be color-selective, it is important to note that most V4 cells sensitive to stimulus wavelength are sensitive to stimulus form as well.

**RECEPTIVE FIELD AND SURROUND INTERACTIONS**

Our results indicate that like V1, area V4 is not specialized for one stimulus quality but rather processes several in parallel. However, we are still left with the problem of what V4 contributes to vision beyond that accomplished in striate cortex. One possible answer to this question may come from the study of the large silent suppressive surrounds that we discovered lie beyond the excitatory receptive fields of most V4 cells. Stimuli placed outside of the excitatory receptive field are without effect themselves, yet are able to suppress the response to a receptive field stimulus (Schein et al., 1983; Desimone and Schein, 1983). These suppressive surrounds can be 30° or more in diameter, extending across the vertical meridian up to 16° or more into the ipsilateral visual field (Moran et al., 1983).

Our experiments with white light revealed that stimuli flashed outside the excitatory receptive fields of V4 cells could have a powerful suppressive effect on the response to a flashed receptive field stimulus. Might this stimulus interaction be spectrally specific? To answer this question, we first determined the optimal wavelength for the receptive field stimulus of a given cell. This optimal colored stimulus was then flashed within the excitatory field and large annuli of varying colors were simultaneously flashed in the surround outside of, but not continuous with, the receptive field. The responses of most V4 cells were almost totally suppressed when the surround stimulus was the same wavelength as the receptive field stimulus. The suppression progressively diminished as the wavelength of the surround stimulus was shifted away from that of the receptive field stimulus, and occasionally the suppression changed to mild enhancement at far removed wavelengths. The surround stimulus itself elicited no response. Thus, most V4 cells respond best to a receptive field stimulus if there is a spectral difference between the receptive field stimulus and its surround.

Other results suggest that form differences are important for many V4 cells as well. Figure 5 shows an example of a cell that responded better to a grating that was restricted to the receptive field than to a grating that covered both the receptive field and its surround. When we varied the frequency of the receptive field and surround gratings independently, we found effects similar to those we found with color, i.e. maximum suppression when the frequencies of the gratings matched.

**What is the significance of the large suppressive surrounds in V4?** Certainly center/surround type organizations have been described at all levels of the visual pathways, and these presumably serve to enhance contrast. Given the much larger receptive field size in V4, we wonder if this organization might serve more global perceptual mechanisms. One possibility is that the spectral interactions reflect a mechanism for color constancy, crucial for accurate color identification (Land and McCann, 1971; Land, 1977). Color constancy depends on spectral interactions over a wide portion of the retina, and V4 cells are influenced by spectral lights throughout a large retinal area. Indeed, Zeki (1980a, 1983a, 1983b, 1983c)
has recently suggested that some V4 neurons demonstrate color constancy, while neurons in striate cortex do not.

There are a number of significant differences between wavelength-selective neurons in V4 and striate cortex. Although both neurons in V4 and double-color-opponent (DCO) neurons in striate cortex (Michael, 1978a, 1978b, 1978c, 1979) respond best to color contrast, V4 cells respond to white light while DCO cells do not; thus, V4 cells are more like the color “filters” predicted by Land’s retinex theory (Land and McCann, 1971; Land, 1977). Another difference is that many striate DCO cells give excitatory responses to stimuli in their surrounds, while V4 cells do not. Finally, our preliminary evidence indicates that the suppressive surrounds of receptive fields in V4 are much larger than surrounds in striate cortex (Moran et al., 1983; and unpublished data). Any or all of these differences might underlie any special role of V4 neurons in color constancy.

In addition to color constancy, the sensitivity of V4 cells to both form and color differences might be useful for figure/ground separation, or “breaking camouflage”, an essential element of both form and color vision. Miezen et al. (1982; and Allman, personal communication) have speculated that some type of figure/ground mechanism may operate in other extrastriate visual areas as well. They have found that for neurons in areas MT and V2 that respond to one direction of stimulus motion within their excitatory receptive field, the response is suppressed by the same motion in the receptive field surround. Thus, these areas may participate in figure/ground separation based on differences in motion. Similar results have recently been reported for neurons in the lateral suprasylvian area of the cat (Rizzolatti and Camarda, 1977; von Grunau and Frost, 1983).

A recent study has reported that neurons in area V2, but not V1, respond to illusory contours even when the stimuli that evoke the illusory contours are located outside of the excitatory receptive field (Von der Heydt et al., 1984). These results, in conjunction with those of our own and others in V2, V4, MT, and the lateral suprasylvian area, suggest that as one moves from striate through prestriate cortex, visual processing by single neurons is influenced by an increasingly wide portion of the visual field. As expected from the anatomy, this trend is continued as one moves from prestripate into the inferior temporal cortex.

SHAPE ANALYSIS IN INFERIOR TEMPORAL CORTEX

Numerous studies have shown that neurons in IT, like neurons in V1, V2, V3 and V4, are sensitive to a variety of stimulus features, both spectral and
Beyond the striate cortex

spatial (Gross et al., 1972; Ridley and Ettlinger, 1973; Jarvis and Mishkin, 1977; Rolls et al., 1977; Mikami and Kubota, 1980; Sato et al., 1980; Fuster and Jervey, 1982; Richmond and Optican, 1983; Desimone et al., 1984). Yet, when IT neurons are studied with bars and edges of different orientations and colors (Gross et al., 1972; Desimone et al., 1984) or sinusoidal gratings varying in spatial frequency (Polien et al., 1984; Desimone, unpublished data), they are no more narrowly tuned along these stimulus dimensions than are cells in striate cortex or V4, and they are generally less so. Given our results in V4, we might expect that the contributions of IT neurons to visual processing lie in more global visual analyses.

The spatial and spectral properties of V4 cells would seem to be useful for extracting object contours and describing object surfaces. However, because V4 cells have comparatively small excitatory receptive fields, an individual V4 cell can provide little information about overall surface shape. By contrast, neurons in IT have large receptive fields. While an excitatory receptive field in V4 that includes the center of gaze is often less than 1\,\text{deg}^2, almost all receptive fields in IT include the center of gaze and average over 625\,\text{deg}^2 (Desimone and Gross, 1979). Thus, IT might synthesize the locally filtered representations conveyed by V4 and provide a more global representation of surface contours, or shape. Evidence has recently been reported for such a global shape-encoding mechanism in IT, a mechanism that is relatively independent of the size and position of the stimulus on the retina (Schwartz et al., 1983).

The shapes used as stimuli in the latter study were based on the Fourier descriptor method of boundary shape analysis (Zahn and Roskies, 1972). According to this method, the local boundary orientation function of a surface 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). Biologically, shape descriptors such as the FDs might be extracted by neural "filters" sensitive to global boundary curvature (Schwartz, 1980). Whether or not IT neurons actually encode shapes using a mechanism similar to the Fourier descriptors, the FDs can provide a tool for probing shape selectivity much the same way that sine-wave gratings can be used to probe spatial selectivity. To explore the shape selectivity of IT neurons, a set of stimuli was created from a set of individual FDs. These shapes varied systematically in their boundary curvature. A few of the shapes used in the study are illustrated in Fig. 6.

About half the visually responsive cells tested in IT were tuned to the frequency of the FD stimuli, much the same way that neurons in V4 or striate cortex are tuned to the spatial frequency of a sinusoidal grating. Cells in striate and prestriate cortex have relatively small, visuotopically organized receptive fields, and thus their response is highly dependent on the size and position of the stimulus. In IT, however, the neurons have very large receptive fields that almost always include the center of gaze, and their selectivity for an FD stimulus was found to be relatively independent of the exact size and position of the stimulus within the receptive field. An example is shown in Fig. 7, which shows the responses of an IT neuron to a series of FD stimuli varying in size. Although the absolute response of the neuron varied with changes in size, and hence did not exhibit size constancy, the preferred FD frequency remained the same. Similar results from changes in size or position were found for over half the IT neurons selective for shape. Thus, the sensitivity of most IT neurons to shape appears to be based on a global property of the shape rather than on the size or location of local contours. Since IT neurons appear to maintain their selectivity for textures and colors throughout a large receptive field as well (Gross et al., 1972; Desimone et al., 1984), they may contribute to mechanisms for the perceptual constancy of objects over changes in size or position.

---

**Fig. 6.** Examples of FD stimuli, which vary in frequency from 2 to 64 cycles per perimeter. Although stimuli are shown in black, the stimuli used in the experiment were normally white on a neutral background.

**Fig. 7.** Responses of an IT neuron to FD stimuli of different sizes. All stimuli were confined to the receptive field and were centered on the fovea. Responses calculated as in Fig. 4. Although the absolute response of the neuron varied with stimulus size, the preferred frequency remained 4 cycles per perimeter.


CONCLUSION

The visual recognition of objects depends on the flow of information along a pathway from striate cortex into the temporal lobe. This pathway is organized primarily as a serial hierarchy, in which processing in areas along the route is multidimensional and increasingly complex at each stage.

The special contribution of neurons in two of these areas, area V4 and the inferior temporal cortex, is not evident from the analysis of their response to simple, local visual features such as orientation or wavelength. In V4, neurons with a relatively small excitatory receptive field are influenced nonetheless by colors and patterns throughout a large portion of the visual field. In inferior temporal cortex, neurons are sensitive to the overall shape of a stimulus, anywhere within a large receptive field. These results suggest that the contributions of extrastriate neurons to pattern recognition may lie in global, rather than local, visual analyses.

Acknowledgements — We gratefully acknowledge the contributions of our collaborators: Thomas D. Albright, Francisco de Monasterio, Charles G. Gross, Mortimer Mishkin, and Eric L. Schwartz.

REFERENCES


Iwai E. and Mishkin M. (1968) Two visual foci in the temporal lobe of monkeys. In Neurophysiological Basis of Learning and Behavior (Edited by Yoshii N. and


Rockland K. J. and Pandya D. N. (1979) Laminar origins and terminations of cortical connections of the occipital lobe in the rhesus monkey. Brain Res. 179, 3-20.


Wong-Riley M. (1978) Reciprocal connections between striate and prestriate cortex in squirrel monkey as demonstrated by combined peroxidase histochemistry and autoradiography. Brain Res. 147, 159-164.


