European Journal of Neuroscience, Vol. 22, pp. 212–224, 2005

Tuning for shape dimensions in macaque inferior temporal cortex

Greet Kayaert,[1] Irving Biederman,[2] Hans P. Op de Beeck[1] and Rufin Vogels[1]
[1]Laboratory Neuro- en Psychofysiologie, K.U. Leuven Medical School, Leuven, Belgium
[2]Department of Psychology and Neuroscience Program, University of Southern California, Los Angeles, CA, USA

Keywords: inferotemporal, object recognition, separable tuning, ventral stream

Abstract

It is widely assumed that distributed bell-shaped tuning (e.g. Radial Basis functions) characterizes the shape selectivity of macaque inferior temporal (IT) neurons, analogous to the orientation or spatial frequency tuning found in early visual cortex. Demonstrating such tuning properties requires testing the responses of neurons for different values along dimensions of shape. We recorded the responses of single macaque IT neurons to variations of a rectangle and a triangle along simple shape dimensions, such as taper and axis curvature. The neurons showed systematic response modulation along these dimensions, with the greatest response, on average, to the highest values on the dimensions, e.g. to the most curved shapes. Within the range of values tested, the response functions were monotonic rather than bell-shaped. Multi-dimensional scaling of the neural responses showed that these simple shape dimensions were coded orthogonally by IT neurons: the degree and direction of responses modulation (i.e. the increase or decrease of responses along a dimension) was independent for the different dimensions. Furthermore, for combinations of curvature-related and other simple shape dimensions, the joint tuning was separable, that is well predicted by the product of the tuning for each of the dimensions. The independence of dimensional tuning may provide the neural basis for the independence of psychophysical judgements of multidimensional stimuli.

Introduction

Visual neurons are believed to code stimulus attributes by overlapping bell-shaped tuning functions with different optima (Poggio & Bizzi, 2004). The tuning of neurons in early visual areas, e.g. for orientation and spatial frequency (De Valois & De Valois, 1990), supports this coding scheme. It is assumed that distributed bell-shaped tuning (e.g. Radial Basis functions; Edelman, 1999; Riesenhuber & Poggio, 2002) also characterizes the shape representation in inferior temporal (IT) cortex, an area that codes for object properties (Logothetis & Sheinberg, 1996; Tanaka, 1996). Demonstrating such tuning requires testing the responses of neurons for different values along dimensions of shape. Using parameterized shapes, Op de Beeck *et al*. (2001) and Kayaert *et al*. (2003) found a systematic relationship between shape similarity and IT response modulation. However, Op de Beeck *et al*. (2001) used few shapes and complex dimensions, and Kayaert *et al*. (2003) searched for responsive neurons with one extreme of the range of parametric variations, which produces a bias in the measured tuning. These limitations of the former studies preclude firm conclusions regarding the characteristics of the tuning along shape dimensions.

Here we examine the tuning of IT neurons to shapes that vary along simple shape dimensions. As in Kayaert *et al*. (2003), the shape dimensions were based on the generalized cone formalization (Marr, 1982; Nevatia, 1982; Biederman, 1987), in which a cross-section, which can vary in size, is swept along a straight or curved axis. In the basic shape set (Fig. 1), a rectangle or triangle was varied along the following dimensions: expansion of the cross-section (taper, for the rectangle only), curvature of the main axis, and positive or negative curvature of the sides. However, in a departure from Kayaert *et al*. (2003) biased sampling, responsive neurons were searched with all stimuli allowing an unbiased measurement of the tuning. Our primary focus was on whether the tuning to these shape dimensions would be bell-shaped.

Human psychophysical studies suggest that curvature and tapering vs. aspect ratio are coded independently (Arguin & Saumier, 2000; Stankiewicz, 2002; Op de Beeck *et al*., 2003). That is, the discrimination along one such dimension is independent of the value on another dimension. Also, the independence of these dimensions is assumed by the Recognition By Components (RBC) theory of object recognition (Hummel & Biederman, 1992). The separability of dimensions at the behavioral level can have several neural correlates. Firstly, neurons tuned for one dimension may be only minimally modulated by changes along another shape dimension. Secondly, neurons may show a joint separable tuning for several shape dimensions. Tuning is termed 'separable' when it is the product of the tuning for each of the dimensions. Such separable tuning is present in V1 for orientation and spatial frequency (Mazer *et al*., 2002) and motion and direction tuning (Grunewald & Skoumbourdis, 2004). Thus, as a second goal we measured the tuning for one dimension as a function of the value on the other dimension, using the shapes of Fig. 2, in which two dimensions were crossed.

Materials and methods

Subjects and recordings

Extracellular single-unit recordings were made in IT of three rhesus monkeys (*Macaca mulatta*) using published techniques (Kayaert *et al*., 2003). During the course of the recordings, we obtained anatomical MRI scans and CT scans with the guiding tube *in situ* in

Correspondence: Dr R. Vogels, as above.
E-mail: rufin.vogels@med.kuleuven.ac.be

Received 23 December 2004, revised 26 April 2005, accepted 1 May 2005

doi:10.1111/j.1460-9568.2005.04202.x

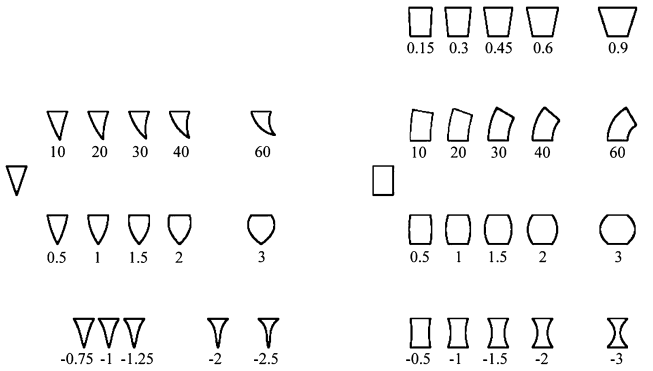


FIG. 1. Stimulus set consisting of variations of a triangle (left) and rectangle (right) on, respectively, three and four dimensions. The dimensions are (from top to bottom) taper, axis curvature, and positive and negative curvature of the sides. The horizontal distances between the stimuli correspond to the magnitudes of the difference between them (within the dimensions). Numbers below the shapes indicate 3D Studio values on the dimension (see Materials and methods).

each monkey. This, together with depth ratings of the white and grey matter transitions and of the skull basis during the recordings, allowed reconstruction of the recording positions. Eye position was measured using either the scleral search coil (two animals) or cornea reflection imaging (ISCAN).

All surgical procedures and animal care were in accordance with the guidelines of the K.U. Leuven Medical School and the European Communities Council Directive 86/609/EEC. The experiments were approved by the Ethical Committee Animal Experiments of the K.U. Leuven. The surgeries (implantation of head post, search coil and recording chamber) were performed under sterile conditions and deep gas (1.2 MAC isoflurane; 50% N_2O/50% O_2) anesthesia. Postoperatively, an analgetic (Tramadol Hydrochloride 15 mg/day) was provided for 48–72 h.

Trials started with the onset of a small square at the display's center on which the monkey was required to fixate. After a fixation period of 300 ms, the fixation target was replaced by a stimulus for 200 ms. If the monkey's gaze remained within a 1.5–2 ° fixation window, the stimulus was replaced again by the fixation spot. Maintaining fixation for at least another 100 ms after stimulus offset was rewarded with a drop of apple juice. All shapes were equally associated with the reward, so that presentation of the reward could not have differentially affected responses to the different stimuli. When the monkey failed to maintain fixation, the trial was aborted and the stimulus was presented during one of the subsequent fixation periods. As long as the monkey was fixating, stimuli were presented with an interval of approximately 1 s. Fixation breaks were followed by a 1 s time-out period before the fixation target was shown again.

For 102 responsive neurons, the responses to the 37 shapes in the stimulus set of Fig. 1 were measured, using interleaved presentations. The presentation of the stimuli was pseudorandom: one of the 37 images was randomly chosen but could not be presented again until all 37 stimuli had appeared once, after which the cycle was repeated. For 62 of these neurons, the 82 unique shapes of the combined stimulus sets of Figs 1 and 2 were tested, also with pseudorandomly interleaved presentations of the different shapes.

Stimuli

The stimuli were white shape outlines (extension approximately 3 °; line width = approximately 0.1 °, luminance approximately 20 candelas/m^2). The shapes were rendered by 3D Studio MAX, release 2.5 [by Autodesk (Kinetix), San Rafael, California, USA] and the edges were extracted by the 'find edges' filter of Adobe Photoshop release 5.5 (by Adobe Systems, San Jose, California, USA), after which the image was inverted (replacing dark parts by light parts and vice versa) to produce white shape outlines on a black background. We used 2D outlines instead of filled shapes to reduce the contribution of luminance/shading variations to response modulation.

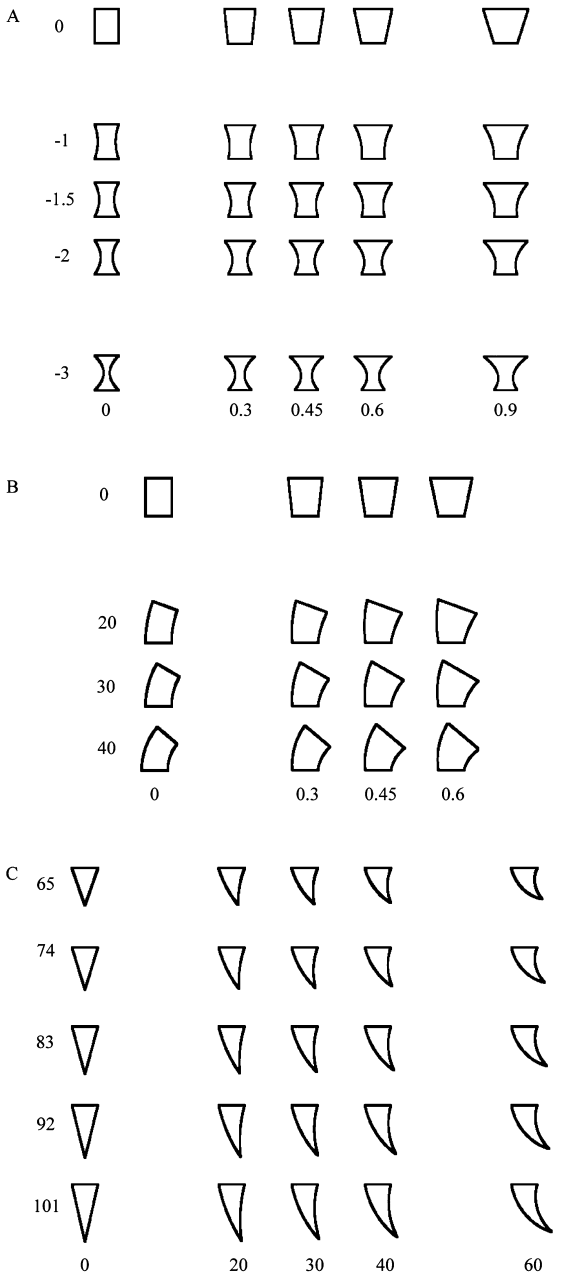


FIG. 2. Stimulus set consisting of shapes varying on two dimensions simultaneously. The shapes vary on negative curvature of the sides (vertical) and taper (horizontal) of the rectangle (A), axis curvature (vertical) and taper (horizontal) of the rectangle (B), axis curvature (horizontal) and aspect ratio (vertical) of the triangle (C). The distances between the stimuli on the horizontal and vertical dimension correspond to the magnitudes of the difference between the shapes on these dimensions. Numbers indicate values on the dimension.

The stimuli are shown in Figs 1 and 2. The shapes were presented on a black background centrally during fixation. They were made by creating a triangle and a rectangle in 3D Studio Max (or actually, the front side of a box and a pyramid) and modifying these along different dimensions.

For the rectangle (see Fig. 1), the dimensions were Axis Curvature (the 'Bend' function in 3D-Studio), Positive and Negative curvature of the sides (using the 'Curve' subfunction of the Taper function in 3D-Studio) and Taper (i.e. the upper side of the rectangle became wider). The amount of modification for each dimension was increased in equal steps, with the exception of the most extreme (rightmost in Fig. 1) stimulus along a dimension, which differed twice as much from the next to last. In all the figures (except Figs 4 and 6), the horizontal distance of the pictured stimuli within the dimensions corresponds with the magnitude of the parametric difference between them. These distances between the stimuli in 3D-Studio (see Fig. 1 for values) corresponded exactly with the distances between direct measurements of the image properties [i.e. length of the upper segment for taper, the angle of the upper segment (i.e. from 0° till −60°) for Curvature of the main axis and the perpendicular distance between the point of maximum curvature and a line drawn from the upper to the lower right corner for Negative and Positive Curvature (a curvature measure called 'Sag', see Foster *et al*., 1993 for a discussion of its relation with human curvature perception)]. The relation between the 3D-Studio values and these image measurements did not change when two modifiers were applied simultaneously (i.e. bend and taper or negative curvature and taper), indicating that the modifiers were applied independently. The distances between the shapes varying on two dimensions simultaneously are indicated by the distances between these stimuli (on the horizontal and vertical dimension) in Fig. 2.

For the triangle the same 3D-Studio modifiers were used to create parametric variations on the dimensions of Axis Curvature and Negative and Positive Curvature of the sides. The distances between the stimuli are the same as for the rectangle variations, except on the Negative Curvature dimension where the first and third stimulus are closer to the second one (by half their original distance), and the most extreme (rightmost) stimulus is closer to the fourth (also by half the original distance), see Fig. 1 for values. The length of the triangle was manipulated by changing the height of the corresponding pyramid in 3D-Studio, see Fig. 2C.

For the shape variations of the rectangle and triangle, the distances between the 3D-Studio values and the values derived from the image measurements corresponded with the pixel-based, physical distances between the stimuli. The latter pixel-based distance between two shapes was computed by filling the interior of the shapes (the same gray-level as the outlines) and computing the Euclidean distance between the gray-levels of the pixels of the images.

Data analysis

The response was defined as the number of spikes, averaged across trials. The median number of trials per stimulus was 13, the lower quartile was 10 and the upper quartile was 20. For the 98 significantly modulating neurons used in the analyses of the responses to the stimuli in Fig. 1, the median number of repetitions of these stimuli was 14, the lower quartile 10 and the upper quartile 20. For the neurons used in the analyses of the three groups of stimuli in Fig. 2, the median numbers of repetitions were, respectively, 13, 13 and 14 trials, the lower quartiles were, respectively, 10, 10 and 11 trials, and the upper quartiles were, respectively, 16, 19 and 19 trials.

The response was measured in a 300-ms interval that started 40–100 ms after stimulus onset (on average 70 ms). The starting point of the interval differed between neurons and depended on the response latency of a particular neuron, which was visually estimated by looking at the peri-stimulus time histograms (PSTH). The onset of the analysis window was chosen conservatively, well before response onset of the neuron. The same analysis window was used for each shape of a particular single neuron. The population PSTHs in Fig. S1 of the Supplementary material show that our analysis windows were tailored to capture the neurons' entire response. Indeed, typically the response of IT neurons lasts longer than the stimulus duration, explaining the 300 ms analysis window. Also note that the population PSTHs show no indication of an effect of reward delivery (starting at 300 ms after stimulus onset).

The significance of the responses modulations to the different shapes was tested by one-way analysis of variance (ANOVA, $P < 0.05$). In all analyses, neurons were only included when there was a significant difference in their responses to the different shapes. When the responses were averaged over neurons, the standard errors of these responses, as shown in Figs 3B, 7B, 8B and 9B, were computed after correcting for the difference in mean response strength among the different neurons, as appropriate for within-subjects (neuron) designs (Loftus & Masson, 1994).

To assess whether the response, averaged across neurons, is significantly modulated by the value along a dimension, we performed a repeated-measure one-way ANOVA on the average responses for the different values along a dimension and this for each of the dimensions, excluding the rectangle and the triangle. To assess whether the population response increased (or decreased) as a function of increasing value along a dimension, we performed contrast analyses (*a priori* planned comparisons) using the following contrasts: for each of the dimensions of the shape variations of the rectangle and for axis curvature and positive curvature of the sides of the triangle, contrast values were ranked according to increasing value along a dimension (again excluding the rectangle or triangle), −55, −30, −5, 20, 70. The contrasts for negative curvature of the sides of the triangle were −3, −2, −1, 2, 4.

To determine how the population of IT neurons represented the shape variations along the different dimensions, we applied standard non-metric multidimensional scaling (MDS; Kruskal & Wish, 1978). The MDS was performed on the Euclidean distances in the high-dimensional neural space (neural distance) for all possible pairs of shape variations in Fig. 1 of either the rectangle or the triangle. Neural distance $= \{[\Sigma^n_i(R_i^1 - R_i^2)^2]/n\}^{1/2}$, with R_i^1 the response of neuron i, averaged over trials, to stimulus 1 and n the number of neurons tested with that pair of stimuli. The dimensionality of the configurations computed by the MDS was chosen using Screen Plot analysis. The 3D and 4D solutions each explained 99% of the variance of the neural distances for the triangular and rectangular shapes, respectively. The low-dimensional configurations were transformed (Procrustes, orthogonal rotation) towards a parametric space in which the different dimensions were orthogonal and distances between the shapes were based on the parametric value along that dimension. To quantify the orthogonality of the neural representation of the dimensions, we fitted straight lines along the position of the shapes of a particular variation in the two-dimensional projections of the rotated 3D or 4D spaces using the Principal Axis method (minimal least-square perpendicular error criterion; Ballard & Brown, 1982).

The responses to the shapes varying on 2D (Fig. 2) were analysed with a 2-way ANOVA in which each factor corresponded to a dimension. In addition to determining the significance of the main and interaction effects, we also computed the strength of association index

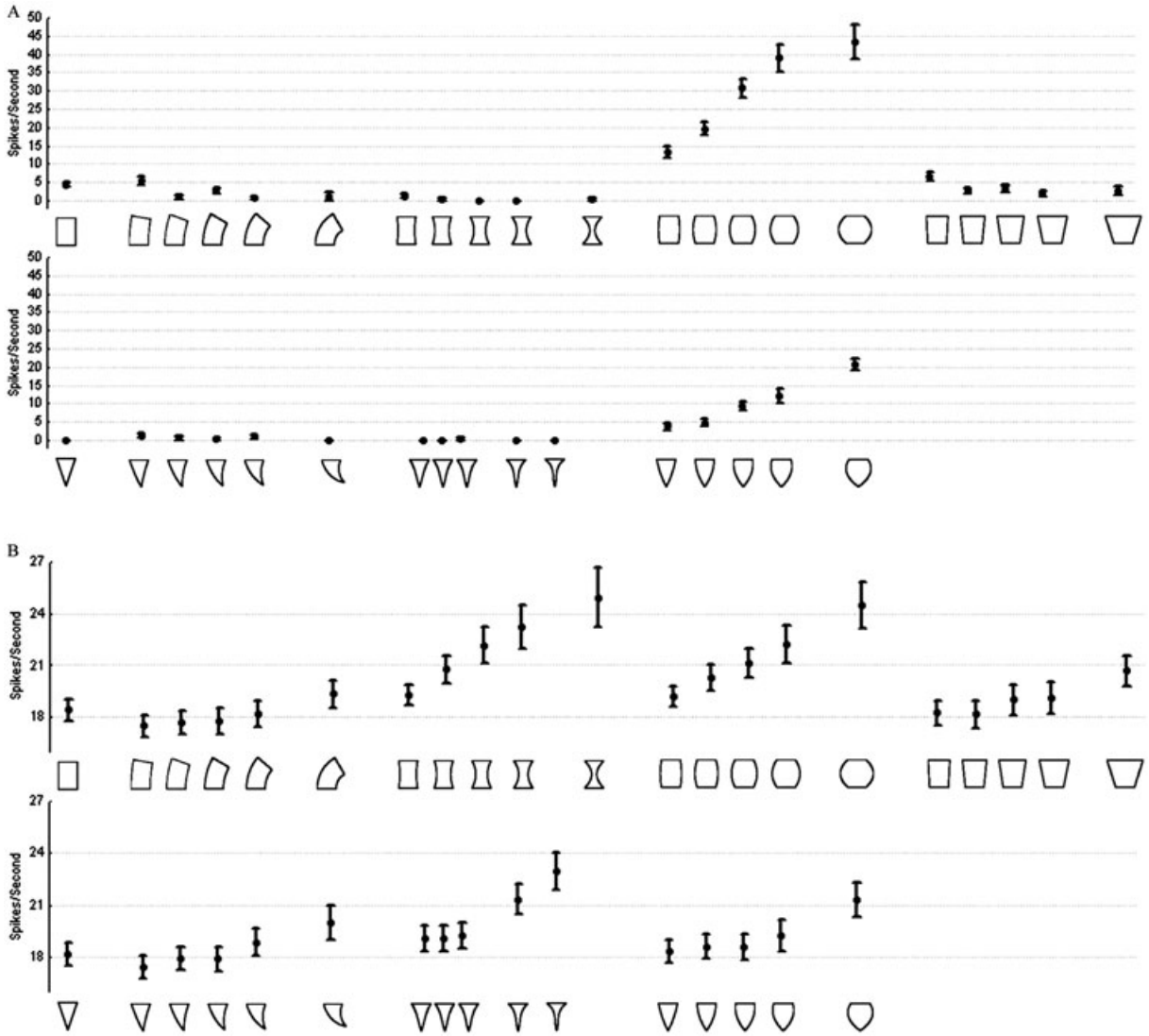


FIG. 3. (A) Responses of a sample neuron. (B) Average responses of 98 neurons that were sensitive to these stimulus variations. The bars represent standard errors.

ω^2 (Kirk, 1968; Erickson *et al*., 2000) for each effect. The latter index takes into account both differences in mean response across conditions as well as within-condition response variations, i.e. trial-to-trial response variability. We used this index as a quantitative measure of the degree of response modulation produced by variations along one dimension (main effects ω^2) or the combination (interaction effect ω^2) of two dimensions.

To assess separability of a pair of dimensions (shapes in Fig. 2), 1D tuning along each of the dimensions was estimated by summing over the other dimension (Mazer *et al*., 2002). The 2D matrix product of the vectors containing these marginal sums corresponds to the predicted 2D separable tuning. Separability was quantified for each neuron by computing the squared Pearson correlation between the measured and predicted separable tuning.

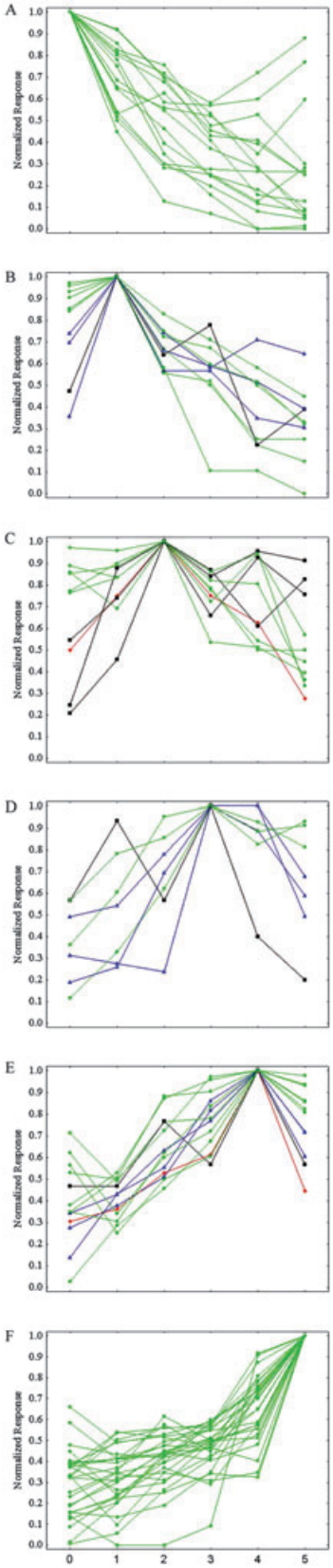
The MDS and the ANOVAs were performed in Statistica (release 6.0, Statsoft, Tulsa, Oklahoma, USA), the Procrustes Rotation was done with PROTEST (created by P.R. Peres-Neto and D. Jackson; freely available at http://www.zoo.utoronto.ca/jackson/pro1.html), and the other major analyses were performed in Matlab house code (Matlab release 5.3., The Mathworks, Natick, Massachusetts, USA).

Results

We recorded from 102 responsive neurons in anterior IT of three monkeys (monkey 1: 37; monkey 2: 20; monkey 3: 45 neurons). These neurons responded to at least one, and usually several, of the shape outlines of the stimulus sets. Across animals, the estimated range of anterior–posterior recording positions of the responsive neurons was between 12 and 21 mm anterior to the external auditory meatus and between 19 and 24 mm lateral to the midline. Most responsive neurons (64%) were recorded from the approximately 3-mm-wide posterior part of the anterior–posterior range (see supplementary Fig. S2). Responsive neurons were located in the lower bank of the superior temporal sulcus (STS) and the cortical convexity lateral to the anterior middle temporal sulcus. Because we found no systematic differences in the tuning to our stimuli among neurons as a function of their anterior–posterior position or between the STS and lateral convexity, we pooled the results of the different recording positions.

First, we will describe the response of all neurons showing significant response modulations when tested with the stimulus set

shown in Fig. 1 (one-way ANOVA, $P < 0.05$, $n = 98$). We will then address the issue of separable coding, using the data from neurons that showed a significant modulation to any of the three stimulus sets of Fig. 2, each with a fully crossed pair of dimensions.



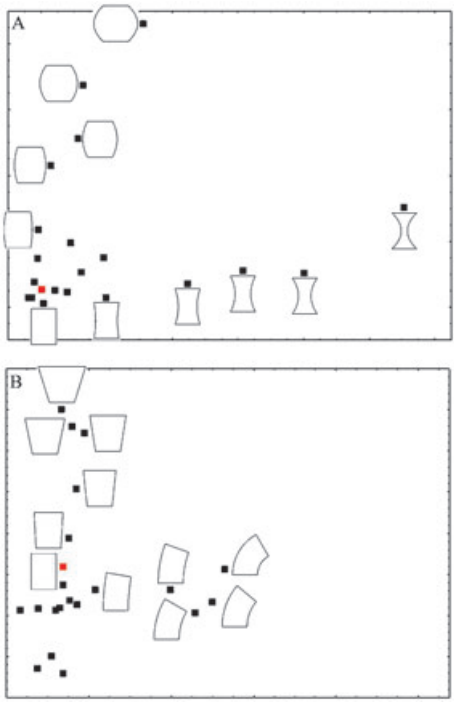


FIG. 5. 2D projections of the 4D MDS solution of the neural distances between the rectangle variations ($n = 98$ sensitive neurons). The dimensions correspond to (A) negative (horizontal) and positive (vertical) curvature of the sides, and (B) axis curvature (horizontal) and taper (vertical). The red square represents the rectangle.

Tuning for shapes varying along simple shape dimensions

The responses of IT neurons were measured to the following unidimensional shape variations shown in Fig. 1: axis curvature and positive (outward) and negative (inward) curvature of the sides of a rectangle and a triangle (Fig. 1). Taper was only manipulated for the rectangle. The responses of a representative neuron to this shape set are shown in Fig. 3A. Two features of the responses of this particular neuron were typical for the entire population of neurons tested. First, the response modulation depended on the dimension: this particular neuron responded most strongly to the outwardly curved shapes. Second, the neuron's strongest response was to the largest value of the dimension; its response increasing monotonically with increasing values along this curvature dimension. Thus, the neuron preferred an extreme value instead of an intermediate value, i.e. it displayed monotonic (Guigon, 2003) instead of a bell-shaped tuning, at least within the range of parametric variations tested.

To summarize the tuning of all neurons (Fig. 4), we selected those dimensions for which the response to the best value of the dimension (including rectangle or triangle) was at least 70% of the overall best response to any stimulus for that neuron, rejecting tuning of which the apparent optimum was evidently non-optimal for the neuron. Second,

FIG. 4. Neural tuning to shape dimensions. (A–F) Each panel shows the selected (see text) tuning functions that peaked at different points along the dimension from the rectangle or triangle (at 0 on the abscissa) in (A) to the most extreme value of the dimension (5 on the abscissa) in (F). Each line represents the normalized responses of one neuron. Colors are chosen in order to facilitate segmentation of the sorts of tuning. The red and blue lines represent the most (only?) convincing cases of bell-shaped tuning. The number of tuning curves (A–F) is, respectively, 16, 10, 10, 7, 13 and 29, which are derived from 8, 7, 9, 7, 13 and 26 different neurons, respectively.

we required that the response modulation – defined as the percent response difference between the maximum response along that dimension and the response to a value that differed by half the full range of that dimension (including rectangle or triangle) – be greater than 40%. This requirement was adopted because neurons that are tuned to the intermediate value can be modulated only along half the range of the dimension, while neurons tuned to an extreme value can show a response difference between the full range of the dimension and thus would have an advantage in passing the 40% response difference criterion. By equating the range over which we tested for modulation for neurons tuned to different values, we thus avoided a bias that would have favored tuning to extreme values. Normalized plots of the tuning functions that met these two criteria are shown in Fig. 4, grouped in terms of the location of the preferred value of a dimension. Note that one neuron can contribute more than one tuning function as it can pass these criteria for more than one stimulus dimension. A total of 54 different neurons and 85 tuning functions were included in the analysis. Figure 4 indicates that some neurons show the strongest response to the triangle or rectangle (Fig. 4A), while others prefer the highest value of the dimension (Fig. 4F), as the neuron shown in Fig. 3. Few convincing examples of tuning for intermediate values were present. Indeed, most of the response profiles with an apparent peak at intermediate values of a dimension were highly asymmetric, suggesting that they are noisy versions of monotonic tuning (only the two red-colored and to a lesser degree the eight blue-colored tuning functions in Fig. 4B–E look convincing). Even if these ambiguous tuning functions are regarded as peaking at intermediate values, the incidence of intermediate tuning is much lower than tuning to the extremes of the dimensions.

Averaging the response of all significantly modulating (one-way ANOVA, $P < 0.05$, $n = 98$) neurons to each of the shapes also shows a stronger response to the highest values on the dimensions (Fig. 3B). The effect of dimension value was significant for each of the dimensions. This was shown with a repeated-measure one-way ANOVA using the responses of the 98 neurons to each of the values of a particular dimension, excluding the rectangle and triangle (as these shapes do not belong to any of the dimensions). The *P*-values of the effect of parametric variation ranged between 0.05 and 0.00001 (median *P*-value 0.00008). *A priori* contrasts defining a linear increase in response as a function of the parametric value on the dimension (see Materials and methods) were significant for each of the seven variations, with *P*-values ranging between 0.04 and 0.0004 (median *P*-value = 0.004). The stronger responses to the highest values of a dimension agree with the higher incidence of monotonic tuning shown in Fig. 4. The increase in response with increasing values of the stimuli on the dimensions is present in the entire time course of the response, as illustrated by supplementary Fig. S1. Similar effects were found when restricting the analysis window from 60 ms to 210 ms after stimulus-onset, i.e. the initial part of the response. Indeed, a repeated-measures ANOVA revealed significant differences in average response using this smaller analysis window, within each of six parametric variations (at least $P < 0.05$ for each dimension), the only exception being the positive curvature of the triangle variation. Thus, as demonstrated in supplementary Fig. S1, the high responsiveness to high values along a dimension is present during stimulus presentation and already in the early part of the response.

Independent coding of simple shape dimensions

At least one model of shape coding (Hummel & Biederman, 1992) assumes that the shape dimensions we have manipulated in the present study are coded independently. If this is the case, one would expect that these dimensions are represented orthogonally in the shape space defined by the neuron's responses. The latter reflects how the neurons represent the shapes and can be analysed by MDS. Each neuron can be conceptualized to correspond to a single dimension of a high-dimensional 'neural' space (with the number of dimensions equalling the number of neurons), and each shape is a point in that space with its coordinates being defined by the responses of the different neurons. MDS attempts to reduce this high-dimensional neural space to a low-dimensional one, keeping the original distances between the shapes in the neural space. If the neurons represent the shape dimensions orthogonally, one expects low-dimensional MDS solutions in which the orthogonal dimensions correspond to the manipulated shape dimensions.

This was tested by performing MDS on the Euclidean neural distances (see Materials and methods) between the three triangle and four rectangle shape variations, for the 98 neurons that significantly modulated to this stimulus set. As expected, MDS produced 3D and 4D solutions which, after Procrustes rotation (see Materials and methods), corresponded to the three and four manipulated shape dimensions, respectively. Note that the three and four dimensionality of the solutions indicates that the neurons responded to the shape of the stimuli and not to possible confounding variables such as size and area, which would have produced 1D solutions. A shape space merely based on pixel differences would have produced, respectively, 2D and 3D solutions for the triangle and rectangle variations (with the dimensions negative and positive curvature of the sides reduced to one). The rank of the shapes in the neural-based configurations correlated perfectly with the ranking in the parametric shape space. The latter is illustrated for two pairs of dimensions in Fig. 5A and B. We fitted straight lines to the points corresponding to the different shapes of the manipulated dimensions in 2D projections of the 4D or 3D configuration. The median angle between these lines for all possible pairs of shape variations was 70 ° (range = 57–97 °; average = 76; $n = 9$), and the median deviation from an angle of 90 ° was 20 ° (range = 4–33; average = 17), indicating a near to complete orthogonality (90 °) of the dimensions.

Interestingly, such near-to-complete orthogonality was also present for the negative and positive curvature dimensions (Fig. 5A). In principle, positive and negative curvature can be two extremes of a single curvature dimension, with zero curvature as the mid-point. However, in the latter case positive and negative curvature would

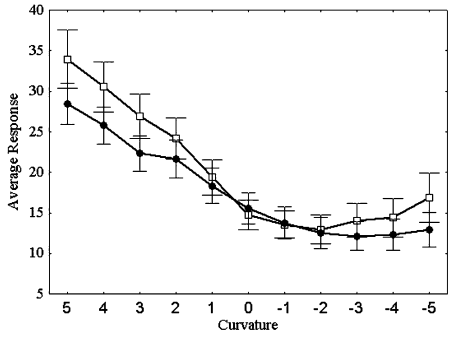


FIG. 6. Mean responses of neurons as a function of curvature sign. Responses of individual neurons were ranked on a negative–positive curvature scale from the variation with the highest value on the preferred dimension, i.e. '5', to the opposite most strongly curved variation, i.e. '−5'. The responses are averaged for the rectangles (empty squares, 56 neurons) and the triangles (full circles, 43 neurons) separately.

correspond to a single instead of the observed two orthogonal MDS dimensions. The MDS result of two orthogonal dimensions reflects the fact that the neurons that responded optimally to the highest value of one of the two dimensions and decreased their response towards either the triangle or the rectangle did not necessarily show a further decrease with increasing values on the other dimension. The tuning to both dimensions, i.e. positive (or outward curvature) and negative (or inward curvature) was essentially independent, so the response could either remain stable, further decrease or even increase again with increasing values on the non-optimal dimension. This is demonstrated in Fig. 6 for those neurons that responded to one of the values of the inward–outward curvature variations by at least 70% of the best response with at least 50% response modulation over the whole range of positive and negative curvature, and did not respond strongest to either the rectangle or the triangle. This selection was done separately for the shapes derived from the rectangle and triangle, respectively. We used 56 neurons for the rectangle variations and 43 for the triangle variations. Then the responses of each individual neuron were ranked on a negative–positive curvature scale from the variation with the highest value on the preferred curvature sign (defined as the curvature sign with the highest average response), to the variation with the highest value on its antipode. Figure 6 shows, in agreement with

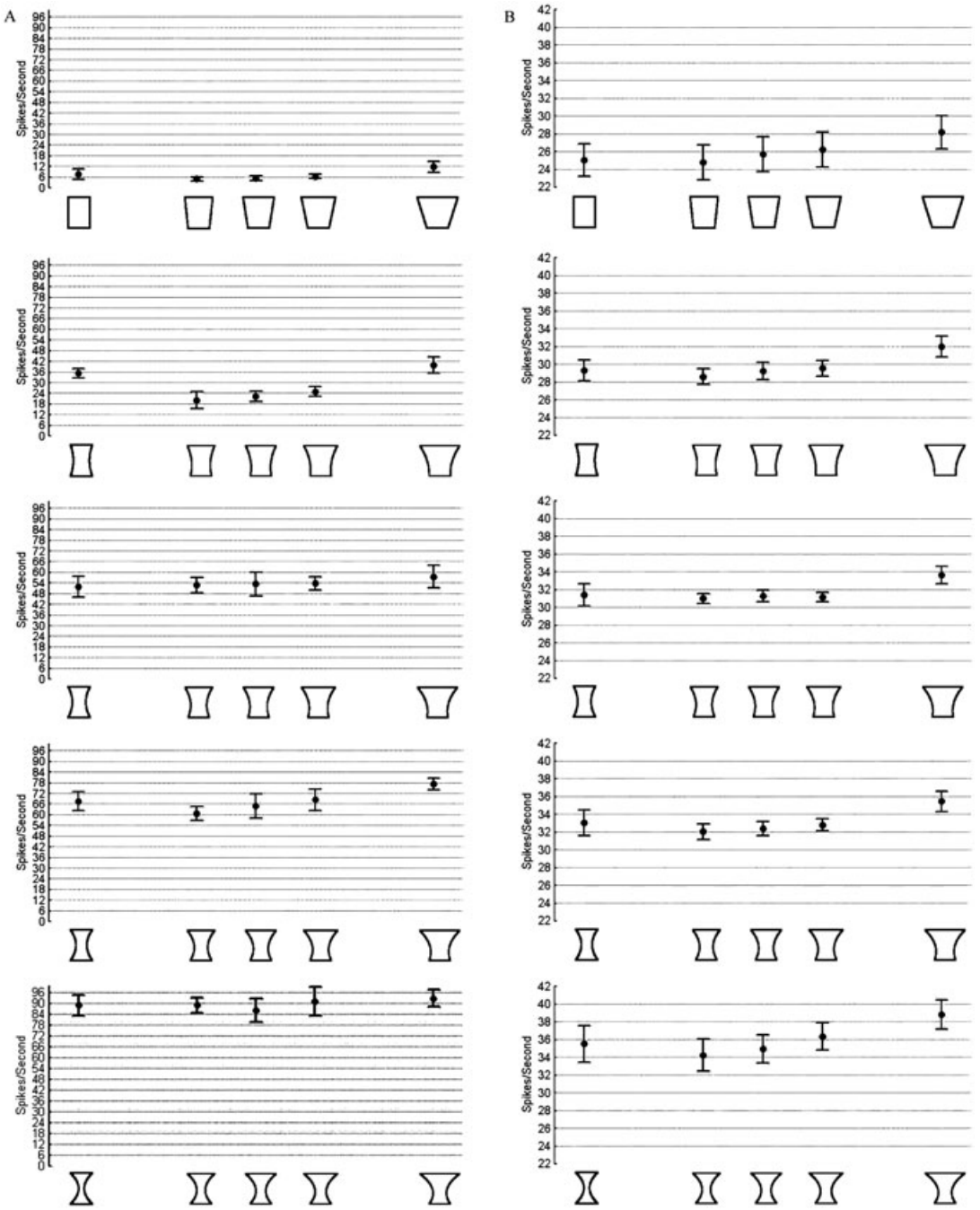


FIG. 7. (A) Response of a sample neuron to the subset of shapes varying on negative curvature of the sides (vertical) and taper (horizontal) of the rectangle. The separability measure (r^2) of the neuron to this shape set was 0.98. (B) Average responses of 54 neurons that were sensitive to these stimulus variations. The bars represent standard errors.

Fig. 4, that the average response decreased monotonically with decreasing degree of optimal curvature but was, on average, less affected by the variations of the opposite curvature sign. Such tuning will produce a 2D MDS solution with positive and negative curvature of the sides as orthogonal dimensions. Thus, IT neurons treat positive and negative curvature of the sides of a shape as two different dimensions and show dimension-specific tuning.

Separability of joint tuning for shape dimensions

So far we have analysed the responses of IT neurons to variations of single, simple shape dimensions. In addition, we combined some of these dimensions to produce three groups of shapes that varied independently on two different dimensions, as shown in Fig. 2. We recorded the responses of 61 neurons to each of these groups of shapes. MDS of the neural distances between the shapes of each group produced 2D configurations. More importantly, the three pairs of dimensions: negative curvature and taper and axis curvature and taper for the rectangle, and axis curvature and aspect ratio for the triangle, were shown to be orthogonal (angles of, respectively, 97, 94 and 95 °) using fitted straight lines in the 2D solutions. Thus, IT neurons treat these dimensions as orthogonal.

Figures 7A, 8A and 9A show the responses of three neurons, which illustrate several features of the population of neurons we tested. First, the degree of modulation depended on the dimension. The first neuron (Fig. 7A) showed the strongest modulation for inward curvature and less for taper. The degree of modulation for a dimension can be quantified by the strength of association index ω^2 (see Materials and methods). For this particular neuron, ω^2 values were 0.76 and 0.01 for the inward curvature and taper dimensions, respectively. The ω^2 for the interaction term was 0. The second neuron (Fig. 8A) shows clear modulation for the two manipulated dimensions (ω^2 for axis curvature = 0.229; ω^2 for taper = 0.57). Again the ω^2 for the interaction of both dimensions was relatively low (0.04). The third neuron (Fig. 9A) responded mostly to variations along the axis

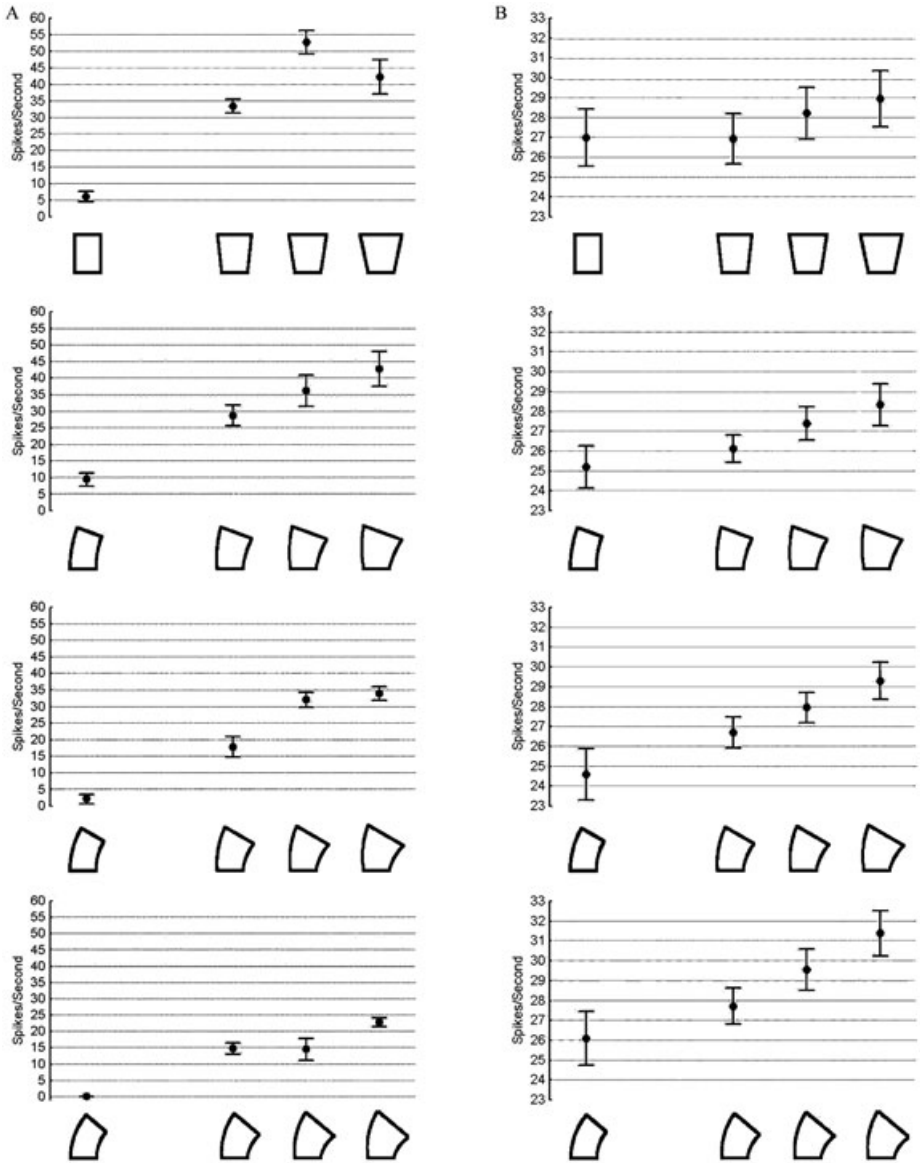


FIG. 8. (A) Response of a sample neuron to the subset of shapes varying on axis curvature (vertical) and taper (horizontal) of the rectangle. The separability measure (r^2) of the neuron to this shape set was 0.95. (B) Average responses of 43 neurons that were sensitive to these stimulus variations. The bars represent standard errors.

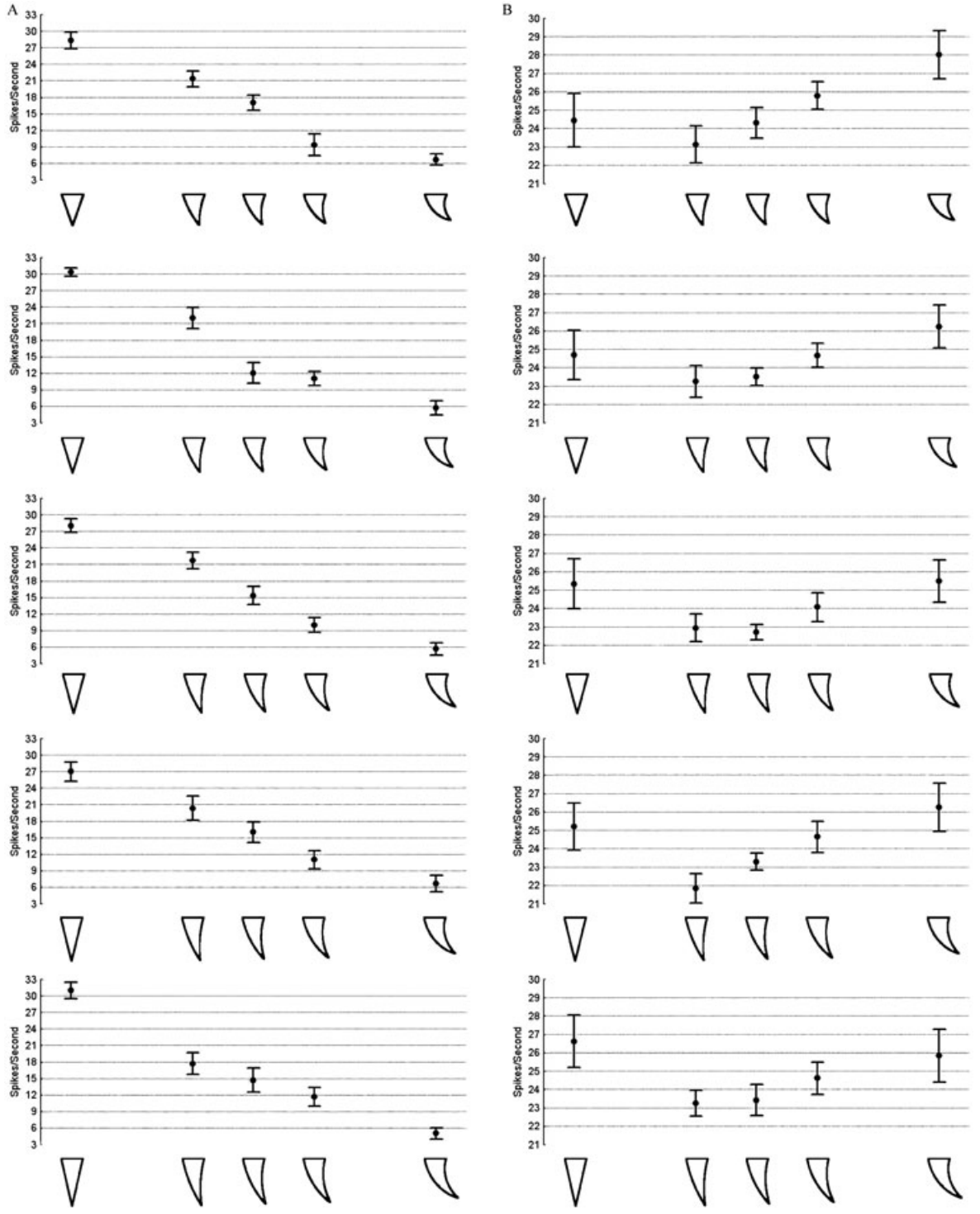


FIG. 9. (A) Response of a sample neuron to the subset of shapes varying on axis curvature (vertical) and length (horizontal) of the triangle. The separability measure (r^2) of the neuron to this shape set was 0.97. (B) Average responses of 41 neurons that were sensitive to these stimulus variations. The bars represent standard errors.

curvature dimension ($\omega^2 = 0.72$), with little modulation for the aspect ratio dimension ($\omega^2 = 0$).

This dimension-dependent tuning can be appreciated by examining Fig. 10, which plots the strength of association values for the pairs of dimensions of each of the three groups. Each point represents a single neuron, and we included only those neurons that modulated significantly as assessed by a one-way ANOVA ($P < 0.05$) with shapes as factor (thus not taking into account the 2D variation). The number of neurons used in this and the following analyses was 54 for the rectangle variations on the dimensions of inward curvature and taper, 43 for the rectangle variations on the dimensions of axis curvature and taper, and 41 for the triangle variations on the dimensions of axis

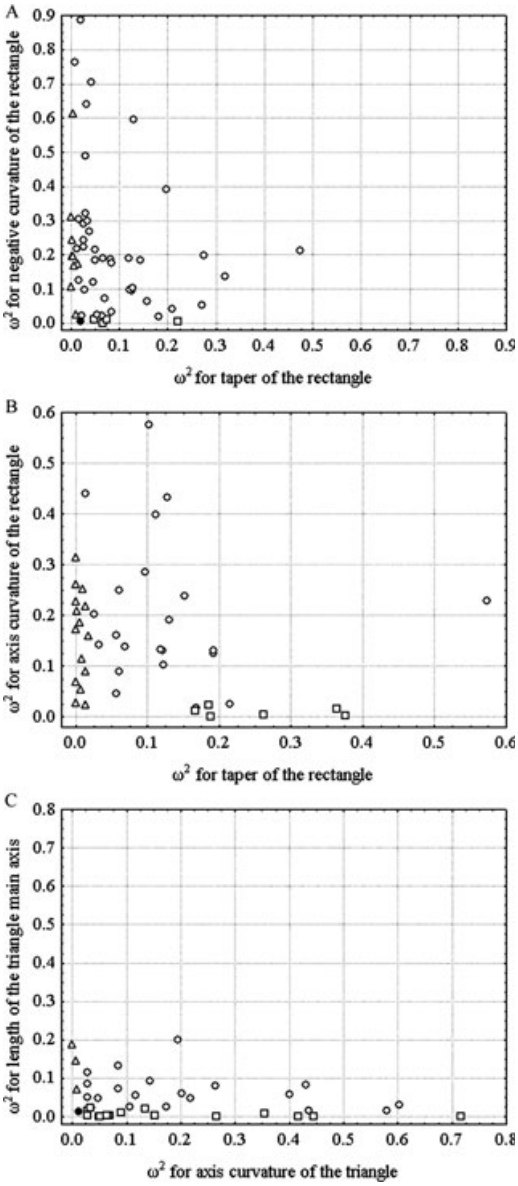


FIG. 10. The strength of association values plotted for the pairs of dimensions of each of the three subsets of shapes varying on two dimensions simultaneously (shapes of Fig. 2). Each point represents a neuron that was sensitive to the shape variations within the particular subset. The open circles represent neurons that significantly modulated to both dimensions, the squares neurons that significantly modulated to the dimension depicted on the abscissa, the triangles neurons that significantly modulated to the dimension depicted on the ordinate, and the filled circles neurons that did not significantly modulate to any of the dimensions (two-way ANOVAs, significant when $P < 0.05$). The total number of neurons is, respectively, 54, 43 and 41 for the three panels.

curvature and aspect ratio. There is a clear clustering of points along the axes of the scatterplots of Fig. 10, indicating that many neurons are more strongly modulated by one than by the other dimension of a pair.

Second, we determined whether the preference of IT neurons for the value of a particular dimension is independent of the value of the shapes along another dimension, i.e. a separable coding of dimensions. The average ω^2 for the interaction of a pair of dimensions was low (0.02 or less for each of the three groups) and much less than the average ω^2 for each dimension, suggesting that there is, at best, only a weak interaction between two dimensions at the neural level. In order

TABLE 1. Separability of dimensions in single neurons

Selection	*n*	Median	First quartile	Third quartile
Axis curvature and taper of the rectangle				
All neurons	43	0.87	0.80	0.92
$\omega^2 > 0.05$	17	0.90	0.88	0.95
$\omega^2 > 0.10$	12	0.92	0.86	0.95
Negative curvature of the sides and taper of the rectangle				
All neurons	54	0.85	0.76	0.92
$\omega^2 > 0.05$	20	0.85	0.81	0.91
$\omega^2 > 0.10$	9	0.89	0.85	0.96
Length and curvature of the main axis of the triangle				
All neurons	41	0.79	0.69	0.90
$\omega^2 > 0.05$	9	0.86	0.80	0.93
$\omega^2 > 0.10$	1	0.74	–	–

n, number of neurons.

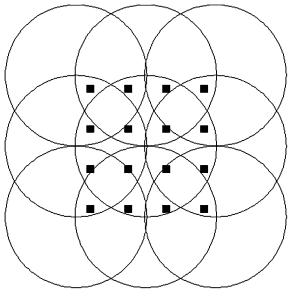


FIG. 11. Schematic representation of a 2D stimulus set (dark squares represent shapes) represented by evenly distributed bell-shaped model neurons (circles).

to quantify the degree of separable coding of these shape dimensions, we determined how well the joint tuning for the 2D shape variations could be predicted from the tuning to the single dimensions (see Materials and methods). Table 1 shows the results of this analysis for the three groups of shape variations. Overall, about 79–87% of the response variance to the different shapes of a group could be explained by the responses to the single dimensions.

This high degree of separability could be a trivial result of neurons being tuned to only one of the two dimensions. To examine this issue, we selected neurons that showed evidence for modulation for both dimensions of a group. Correlating ω^2 and the *P*-values of the two-way ANOVAs showed that ω^2 larger than 0.05 corresponded to highly significant effects with *P*-values smaller than 0.001. The median separability of those neurons for which the ω^2 was larger than 0.05 for both dimensions (or with an interaction ω^2 larger than 0.05) ranged between 0.85 and 0.90 across the three groups (Table 1), which are even somewhat larger values than observed for the whole sample. Using a stricter criterion of ω^2 larger than 0.1 reduced the number of neurons but produced similar results (Table 1). Thus, we conclude that the joint tuning for these simple shape dimensions is separable, i.e. is predicted well by the tuning along each of the two dimensions.

Finally, the strongest response of each of the example neurons was observed at the extreme values of one (Figs 7A and 9A) or two dimensions (Fig. 8A). Monotonic tuning, at least within the manipulated range of the dimensional variations, was typical across the population. Figures 7, 8 and 9B show the mean responses, averaged across neurons, to the shapes for each of the three groups of shapes. It is evident that the average responses were the largest at the extremes of the dimensional variations. The differences in average response

within the stimulus sets were significant (repeated-measure ANOVA, $P < 0.01$ for each of the three stimulus sets). Thus, the stronger responses to the extremes of a single dimension that we found when examining the responses for shape variations along a single dimension (see above; Figs 3B and 4) are also present when two dimensions are varied simultaneously.

Discussion

Using parametric variations of simple shape dimensions such as curvature and taper, we showed that single IT neurons reveal highly systematic response modulations to these shape variations. This result agrees with previous observations by Op de Beeck *et al*. (2001) using more complex shapes that varied along Fourier Boundary Descriptor dimensions, and by Kayaert *et al*. (2003) using 3D shapes or their silhouettes and similar shape variations as in the present study. Unlike the latter study, all neurons in the present study were searched and tested with the same set of parameterized shapes, providing an unbiased measurement of the tuning profiles of IT neurons to these sets of shapes. We found that the tuning profiles of IT neurons for these shape variations were heavily skewed: within the range of the manipulated dimensional variations, the large majority of IT neurons showed monotonic tuning profiles along one or more of these shape dimensions. An analysis of the responses of the population of neurons showed that these simple shape dimensions are represented as (quasi-) orthogonal dimensions. Many neurons were strongly modulated by one dimension but much less so along the other dimension. Furthermore, we showed that the joint tuning for combinations of two of these shape dimensions is highly separable, i.e. can be described as a product of the tuning for each of the dimensions.

Current computational models of shape coding generally employ 'value' or 'local' coding (e.g. radial basis functions) to represent shape (Edelman, 1999; Riesenhuber & Poggio, 2002). Indeed, shapes are believed to be coded by using overlapping bell-shaped tuning functions with different optima in a multidimensional space. Current evidence for such bell-shaped tuning in IT is based on studies showing tuning for the 3D view of objects (Logothetis *et al*., 1995) or faces (Perret *et al*., 1991). However, it should be noted that rotating objects or faces can lead to the sudden (dis)appearance of features, and thus it can be argued that rotation influences not only the values along a dimension but also the stimulus dimension itself. Thus, a critical question is whether the present results using shape variations along well-defined stimulus dimensions agree with this widely accepted coding scheme. Figure 11 shows a 'classical' set of bell-shaped tuning functions that cover a 2D space. The horizontal and vertical dimensions of this figure represent shape variations along two dimensions (e.g. inward curvature and taper of the rectangle), and for the sake of clarity, only three values of preferred neural tuning (i.e. the circles) are shown for each dimension. Bell-shaped tuning, evenly distributed in a multidimensional space, would produce equal responses to all values of the two dimensions when the responses across neurons and dimensions are averaged. However, our data show another pattern: markedly stronger responses to extreme than to intermediate values.

The above argument assumed equal distribution of optima of the overlapping tuning curves. However, the present results are in agreement with a bell-shaped tuning model in which there is an unequal distribution of cell optima for shape, with the majority of neurons being tuned to the extremities of our manipulated range (either the zero or the highest values on the dimensions with, on average, stronger responses to the highest values) or to even more extreme values not present in our stimulus set. Thus, it cannot be excluded that IT neurons show bell-shaped tuning at more extreme values outside of or at the borders of the range that we used, and that the apparent monotonic tuning (Guigon, 2003) that we observed is only one side of such a bell-shaped tuning. But this would imply: (1) that the tuning is relatively broad; and (2) that the optima are located at extreme values of a dimension. The latter reduces the effective response functions within the manipulated range to monotonic functions, with the consequence that differences among these perceptually distinct shapes are, in effect, coded in IT by monotonic functions. We note that for most of the dimensions studied, more extreme values would produce qualitative and not just metric changes in the stimuli. More extreme negative curvature of the sides would produce two shapes touching at a single point that would give the impression of different parts (Hoffman & Richards, 1984; Kimia *et al*., 1995). More extreme positive curvature would change the orientation (i.e. major axis) of the shape from vertical to horizontal. A greater bend in the axis would produce a two-part L-shape, with the parts almost orthogonal to each other or a U-shape in which one leg of the shape is parallel to the other.

Our suggestion that IT neurons show a strong preference towards the extremes of shape dimensions agrees with other recent work on shape tuning in the ventral visual stream. In a series of elegant studies, Connor and co-workers have modeled V4 (Pasupathy & Connor, 2001) and posterior IT (Brincat & Connor, 2004) neuronal responses to a set of curved shapes by bell-shaped functions tuned in curvature × angular position × orientation spaces. Interestingly, the fitted optimal curvatures of V4 neurons were clearly clustered around either extreme or zero curvatures, with a relatively small number of neurons preferring intermediate curvature values (Fig. 7B in Pasupathy & Connor, 2001). Modeling the inputs of posterior IT neurons, to explain responses to a large set of relatively complex shapes, yielded similar strong trends towards optima at extreme curvatures (Brincat & Connor, 2004; Brincat & Connor, personal communication). This agrees with the present results showing clear peaks at extreme curvatures in more anterior IT neurons. In fact, it is possible that the simple shapes used in our study isolated one or a few of the input units of IT neurons modeled by Brincat and Connor. Thus, the tuning functions to the simple shape manipulations that we observed might strongly reflect the tuning properties of neurons at earlier levels, perhaps V4, that provide input to IT.

Another set of results that is in agreement with our finding of IT tuning optima at the extremes of a stimulus dimension was obtained with human faces that varied systematically along 'identity' dimensions in a multidimensional face space (Leopold *et al*., 2003). The center of this space corresponded to the average or prototypical face. Each dimension corresponds to the face of a particular individual, and the distance between the center and a particular value along such an identity dimension reflects the amount of deviation of the face from the average, i.e. the caricature level. Leopold *et al*. (2003) found that most IT neurons showed monotonic response functions along an identity dimension with stronger responses at the extreme caricatures. They interpreted their results in terms of norm-based or prototype-referenced coding in which faces are encoded as a deviation from the average face (Rhodes *et al*., 1987, 1998; Valentine, 1991; Leopold *et al*., 2001). One might interpret our results also in the context of a norm-based model of shape coding, in which shapes are encoded as deviations from a norm in a multidimensional shape space. Because shapes can vary widely, it is likely that several of these 'norms' will be present, producing several 'local' multidimensional shape spaces. In our stimulus set, the rectangle and triangle might represent such norms. However, we observed neurons tuned to the high values of a

dimension, as well as neurons tuned to the 'zero' value of the dimension (but less to intermediate values), which argues against a simple norm-based scheme for shape coding (also see Rhodes *et al*., 1998 for psychophysical evidence against norm-based shape coding). A related coding scheme that agrees with both our shape and the Leopold *et al*. (2003) face data represents shapes with respect to 'anchors', which would be highly distinctive shapes with extreme values along one or more shape dimensions (e.g. zero or a large curvature, or a high taper).

Animals in the Leopold *et al*. (2003) study were extensively trained to categorize the face stimuli according to stimulus identity. However, the monkeys in the present study were not trained to discriminate the stimuli and thus did not have to create a representation of the stimulus space. Nevertheless, they were exposed extensively to the shapes during the recording sessions. Is it possible that the observed tuning resulted from this exposure? It is known that repeated presentations of stimuli can produce an overall reduction in response (Li *et al*., 1993; Ringo, 1996; Ranganath & Rainer, 2003), and this familiarity or adaptation effect may increase with the average similarity between the shape and the other shapes of the set. However, response differences between highly and less familiar stimuli, i.e. effects of presentation frequency, are overall rather small in IT (Erickson *et al*., 2000; Baker *et al*., 2002) and unlikely to explain the full range of response modulations that we obtained in the present study. Another possibility is that IT neurons learned to code for these stimuli by creating monotonic tuning for extreme values of the stimulus set. However, this would be quite a dramatic effect of experience on IT tuning properties. Indeed, given the agreement between the preference for curvature extremes in the V4 data of Pasupathy and Connor and in our data, it is more likely that the monotonic response functions we observed do not result merely from exposure during the experiment and may reflect a default shape-encoding mechanism in the primate ventral stream. Nonetheless, the contribution of adaptation-based mechanisms and/or learning to the observed monotonic tunings should be investigated in future research.

Despite the strong anisotropy in preferred shape and the high incidence of monotonic tuning profiles, MDS showed that these neurons can represent the differences among the shapes quite adequately. The neurons varied in their degree of modulation, and neurons could prefer the higher or smallest value of a dimension. A neuronal population with these two properties can represent the differences among the shapes. Subtracting the activity of neurons responding to the largest value of a dimension (positively sloped function) from the activity of neurons responding to the zero value (negatively sloped) can, in principle, code for small differences among the shapes that vary along a dimension. It has been shown that such a coding strategy using monotonic response functions has the advantage of reducing correlated trial-to-trial variations in response (Romo *et al*., 2003), thus increasing coding accuracy.

In the present study, we deliberately chose shape dimensions – curvature, tapering and aspect ratio – for which human psychophysical studies have provided evidence for independent coding (Arguin & Saumier, 2000; Stankiewicz, 2002; Op de Beeck *et al*., 2003). These studies show that the discrimination along one such dimension (e.g. curvature or tapering) can be performed independently of the value on another dimension (e.g. aspect ratio). The independent, separable tuning that we found renders it relatively straightforward to discriminate values of a dimension (i.e. degree of curvature) independent from variations along another dimension, i.e. by pooling across neurons that prefer the same extreme of a dimension.

An important but difficult to answer question is whether the dimensions we manipulated correspond to (a subset of) the principal shape dimensions used by the visual system to represent shape, as postulated by the RBC theory of object representation (Hummel & Biederman, 1992; Stankiewicz, 2002). In the present study, the strongest modulations were present with curvature-related shape dimensions (inward or outward curvature and axis curvature). Curvature has been shown to be critical when modeling the shape selectivity of V4 and posterior IT neurons (Pasupathy & Connor, 2001, 2002; Brincat & Connor, 2004). Also, recent human psychophysical work by Habak *et al*. (2004) suggests that curvature is an important shape property. The importance of curvature-related shape dimensions for shape encoding might explain why we found good separable coding of curvature dimensions and other dimensions. It should be noted, though, that we did not manipulate just 'curvature' as a single feature: the three curvature dimensions proved to be independent shape manipulations with respect to the tuning of IT neurons. It is currently unknown whether shape dimensions unrelated to curvature manipulations show separability. Also, it is not known the extent to which similar separable behavior is present when changing curvature or other features of more complex shapes than the ones used here, although Hayworth & Biederman (2004) have shown that the aspect ratio of a part and its position with respect to another part of a two-part object are separable features.

In general, the present study indicates that IT neurons show rather systematic behavior to simple parametric variations of relatively simple shapes. Single IT units are tuned preferably to extremes of such simple shape dimensions and show separable encoding for at least combinations of curvature-related and other simple shape dimensions.

Supplementary material

The following supplementary material may be found on:
http://www.blackwellpublishing.com/products/journals/suppmat/EJN4202/EJN4202sm.htm

Fig. S1. Population PSTHs showing the average responses of all 102 responsive neurons to each of the seven shape dimensions used in this study.

Fig. S2. Structural MRI scan of monkey M3 obtained after the recordings.

Acknowledgements

The technical help of M. De Paep, P. Kayenbergh, G. Meulemans, G. Vanparrijs and W. Depuydt is gratefully acknowledged. This study was supported by Human Frontier Science Program Organization RG0035/2000-B (R.V. & I.B.), Geneeskundige Stichting Koningin Elizabeth (R.V.) and James S. McDonnell Foundation 99-53 (I.B). H.P.O. is a postdoctoral fellow of the Fund for Scientific Research (FWO) Flanders.

Abbreviations

IT, inferior temporal; MDS, multidimensional scaling; PSTH, peri-stimulus time histograms; RBC, Recognition By Components; STS, superior temporal sulcus.

References

Arguin, M. & Saumier, D. (2000) Conjunction and linear non-separability effects in shape encoding. *Vis. Res.*, **40**, 3099–3115.

Baker, C.I., Behrmann, M. & Olson, C.R. (2002) Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat. Neurosci.*, **5**, 1210–1216.

Ballard, D.H. & Brown, C.M. (1982) *Computer Vision*. Prentice Hall, Englewood Cliffs, NJ.

Biederman, I. (1987) Recognition-by-components: a theory of human image understanding. *Psychol. Rev.*, **94**, 115–147.
Brincat, S.L. & Connor, C.E. (2004) Underlying principles of visual shape selectivity in posterior inferior temporal cortex. *Nat. Neurosci.*, **7**, 880–886.
De Valois, R.L. & De Valois, K.K. (1990) *Spatial Vision*. Oxford University Press, New York.
Edelman, S. (1999) *Representation and Recognition in Vision*. MIT Press, Cambridge, MA.
Erickson, C.A., Jagadeesh, B. & Desimone, R. (2000) Clustering of perirhinal neurons with similar properties following visual experience in adult monkeys. *Nat. Neurosci.*, **3**, 1143–1148.
Foster, D.H., Simmons, D.R. & Cook, M.J. (1993) The cue for contour-curvature discrimination. *Vis. Res.*, **33**, 329–341.
Grunewald, A. & Skoumbourdis, E.K. (2004) The integration of multiple stimulus features by V1 neurons. *J. Neurosci.*, **24**, 9185–9194.
Guigon, E. (2003) Computing with populations of monotonically tuned neurons. *Neural Comput.*, **15**, 2115–2127.
Habak, C., Wilkinson, F., Zakher, B. & Wilson, H.R. (2004) Curvature population coding for complex shapes in human vision. *Vision Res.*, **44**, 2815–2823.
Hayworth, K.J. & Biederman, I. (2004) Parts and relations are analyzable sources of shape variation: evidence for structural descriptions. *J. Vision*, **3**, 514.
Hoffman, D.D. & Richards, W.A. (1984) Parts of recognition. *Cognition*, **18**, 65–96.
Hummel, J.E. & Biederman, I. (1992) Dynamic binding in a neural network for shape recognition. *Psychol. Rev.*, **99**, 480–517.
Kayaert, G., Biederman, I. & Vogels, R. (2003) Shape tuning in macaque inferior temporal cortex. *J. Neurosci.*, **23**, 3016–3027.
Kimia, B.B., Tannenbaum, A.R. & Zucker, S.W. (1995) Shapes, shocks, and deformations, I: the components of shape and the reaction-diffusion space. *Int. J. Computer Vision*, **15**, 189–224.
Kirk, R.E. (1968) *Experimental Design: Procedures for the Behavioral Sciences.* Brooks/Cole, Belmont, CA.
Kruskal, J.B. & Wish, M. (1978) *Multidimensional Scaling*. Sage Publications, Beverly Hills, CA.
Leopold, D.A., Bondar, I.V., Giese, M.A. & Logothetis, N.K. (2003) Prototype-referenced encoding of faces in the monkey inferotemporal cortex. *Program No. 590.7 2003 (Abstract) Viewer/*Itinerary Planner. Society for Neuroscience, Washington, DC.
Leopold, D.A., O'Toole, A.J., Vetter, T. & Blanz, V. (2001) Prototype-referenced shape encoding revealed by high-level aftereffects. *Nat. Neurosci.*, **4**, 89–94.
Li, L., Miller, E.K. & Desimone, R. (1993) The representation of stimulus familiarity in anterior inferior temporal cortex. *J. Neurophysiol.*, **69**, 1918–1929.
Loftus, G.R. & Masson, M.E.J. (1994) Using confidence intervals in within-subject designs. *Psychonomic Bull. Rev.*, **1**, 476–490.
Logothetis, N.K., Pauls, J. & Poggio, T. (1995) Shape representation in the inferior temporal cortex of monkeys. *Curr. Biol.*, **5**, 552–563.
Logothetis, N.K. & Sheinberg, D.L. (1996) Visual object recognition. *Annu. Rev. Neurosci.*, **19**, 577–621.
Marr, D. (1982) *Vision*. W.H. Freeman, San Francisco, CA.
Mazer, J.A., Vinje, W.E., McDermott, J., Schiller, P.H. & Gallant, J.L. (2002) Spatial frequency and orientation tuning dynamics in area V1. *PNAS*, **99**, 1645–1650.
Nevatia, R. (1982) *Machine Perception*. Prentice Hall, Englewood-Cliffs, NJ.
Op de Beeck, H., Wagemans, J. & Vogels, R. (2001) Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat. Neurosci.*, **4**, 1244–1252.
Op de Beeck, H., Wagemans, J. & Vogels, R. (2003) The effect of category learning on the representation of shape: dimensions can be biased but not differentiated. *J. Exp. Psychol. Gen.*, **132**, 491–511.
Pasupathy, A. & Connor, C.E. (2001) Shape representation in area V4: position-specific tuning for boundary confirmation. *J. Neurophysiol.*, **86**, 2505–2519.
Pasupathy, A. & Connor, C.E. (2002) Population coding of shape in area V4. *Nat. Neurosci.*, **5**, 1332–1338.
Perret, D.I., Oram, M.W., Harries, M.H., Bevan, R., Hietanen, J.K., Benson, P.J. & Thomas, S. (1991) Viewer-centered and object-centered coding of heads in macaque temporal cortex. *Exp. Brain Res.*, **86**, 159–173.
Poggio, T. & Bizzi, E. (2004) Generalization in vision and motor control. *Nature*, **431**, 768–774.
Ranganath, C. & Rainer, G. (2003) Cognitive neuroscience: neural mechanisms for detecting and remembering novel events. *Nat. Rev. Neurosci.*, **4**, 193–202.
Rhodes, G., Brennan, S. & Carey, S. (1987) Identification and ratings of caricatures: implications for mental representations of faces. *Cog. Psychol.*, **19**, 473–497.
Rhodes, G., Carey, S., Byatt, G. & Proffitt, F. (1998) Coding spatial variations in faces and simple shapes: a test of two models. *Vis. Res.*, **38**, 2307–2321.
Riesenhuber, M. & Poggio, T. (2002) Neural mechanisms of object recognition. *Curr. Opin. Neurobiol.*, **12**, 162–168.
Ringo, J.L. (1996) Stimulus specific adaptation in inferior temporal and medial temporal cortex of the monkey. *Behav. Brain Res.*, **76**, 191–197.
Romo, R., Hernández, A., Zainos, A. & Salinas, E. (2003) Correlated neuronal discharges that increase coding efficiency during perceptual discrimination. *Neuron*, **38**, 649–657.
Stankiewicz, B.J. (2002) Empirical evidence for independent dimensions in the visual representation of three-dimensional shape. *J. Exp. Psychol. Hum. Percept. Perf.*, **28**, 913–923.
Tanaka, K. (1996) Inferotemporal cortex and object vision. *Annu. Rev. Neurosci.*, **19**, 109–139.
Valentine, T. (1991) A unified account of the effects of distinctiveness, inversion, and race in face recognition. *Q. J. Exp. Psychol.*, **43A**, 161–204.