Population coding of shape in area V4

Anitha Pasupathy[1] and Charles E. Connor[2,3]

[1] *Center for Learning and Memory and Department of Brain and Cognitive Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

[2] *Department of Neuroscience, Johns Hopkins University School of Medicine, Baltimore, Maryland 21218, USA*

[3] *Zanvyl Krieger Mind/Brain Institute, Johns Hopkins University, Baltimore, Maryland 21218, USA*

Correspondence should be addressed to C.E.C. (connor@jhu.edu)

Published online 11 November 2002; doi:10.1038/nn972

Shape is represented in the visual system by patterns of activity across populations of neurons. We studied the population code for shape in area V4 of macaque monkeys, which is part of the ventral (object-related) pathway in primate visual cortex. We have previously found that many macaque V4 neurons are tuned for the curvature and object-centered position of boundary fragments (such as 'concavity on the right'). Here we tested the hypothesis that populations of such cells represent complete shapes as aggregates of boundary fragments. To estimate the population representation of a given shape, we scaled each cell's tuning peak by its response to that shape, summed across cells and smoothed. The resulting population response surface contained 3–8 peaks that represented major boundary features and could be used to reconstruct (approximately) the original shape. This exemplifies how a multi-peaked neural population response can represent a complex stimulus in terms of its constituent elements.

Shape information is distributed across populations of neurons in the ventral pathway of primate visual cortex[1,2]. The population code for shape has to accommodate the virtual infinity of possible objects as well as the variability of a given object's retinal image. This difficult representational problem could be solved by encoding shapes in terms of their component parts[3–13]. A parts-based coding scheme, like an alphabet, would have the combinatorial power to represent an infinite variety of objects using a finite number of elements. Consistent with this theory, neural tuning for object parts is common in the ventral pathway[14–21]. Here we report on parts-based population coding in macaque V4, a ventral pathway area known to carry information about shape[19,20,22–24], color[25] and texture[26].

Most studies of neural population coding have focused on representation of a single scalar value. A well-known example is the work showing that neural populations in macaque primary motor cortex (M1) encode arm reach direction[27]. In this analysis, each cell was treated as a unit vector oriented at its reach-direction tuning peak. Each vector was scaled by its cell's activity during a given reach. The scaled vector sum closely predicted the actual reach direction in all cases. The vector sum method and other basis function decoding schemes are exact only under certain conditions[28,29], but neural implementation is straightforward and compatible with neural tuning functions in a wide variety of systems. Decoding schemes based on conditional probability distributions of neural response patterns (Bayesian inference, maximum likelihood inference) are more accurate and general, and can also be implemented in biologically realistic networks[30,31]. As we were interested in feasibility rather than optimal coding accuracy in the current study, we chose the basis function approach, which corresponds to the simplest neural decoding mechanism. (Accurate basis function decoding would imply accurate Bayesian or maximum likelihood decoding *a fortiori*.)

Our analysis was based on two-dimensional (2D) Gaussian functions in a curvature × angular position domain. For each V4 neuron in our sample, we determined the Gaussian that best described the curvature and position of boundary fragments (embedded in complete shapes) to which the neuron responded. The population response to a given shape was estimated by weighting each Gaussian peak by the corresponding neuron's response to that shape and then summing across neurons. The weighted sum contained peaks representing the major boundary features of the shape. The accuracy of this representation was confirmed by using the population peak values to reconstruct an approximation to the original shape.

Results

We estimated population-level representations of moderately complex silhouette-type shapes in macaque monkey area V4. We based these estimates on the responses of 109 V4 neurons that showed sensitivity to complex shape in preliminary tests. Boundary shape tuning in such cells can be conveniently (though not necessarily uniquely) characterized in terms of boundary curvature and angular position[20]. These dimensions capture the two critical elements of a parts-based representation: part shape (curvature) and part position (specifically object-centered position[32]). The stimulus set (**Fig. 1a**, white icons) consisted of 49 shapes (constructed by quasi-factorial combination of convex and concave boundary fragments) presented at eight orientations. Each stimulus was flashed for 500 ms at the center of the neuron's receptive field while the monkey performed a fixation task. Response rates were calculated by summing spike occurrences during the 500-ms stimulus presentation period and then subtracting average background rate (derived from null stimulus presentation periods). An example neuron's average responses are indicated by the gray levels of the circular backgrounds sur-

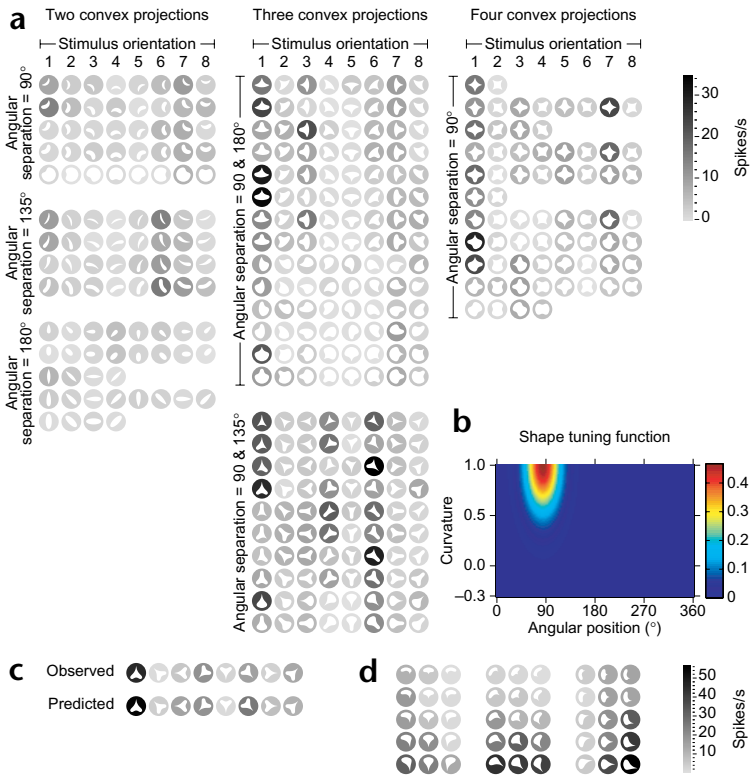

Fig. 1. Single neuron shape-tuning example. (**a**) Responses of an individual V4 neuron are represented by shades of gray surrounding each stimulus icon. The response to each stimulus was averaged across five presentations. The scale bar (right) shows that mean response rates ranged from 0 (light gray) to 34 (dark gray) spikes/s. The stimulus set comprised most of the geometrically feasible combinations of five standard boundary fragments: sharp convex, medium convex, broad convex, broad concave and medium concave curves. Each combination was presented at eight orientations (rows), or fewer if rotational symmetry made some orientations redundant. The stimuli are arranged here into three large blocks (left, middle, right) according to how many convex projections they contained (two, three or four, respectively). They are also blocked in the vertical direction according to the angular separations between convex projections. The stimuli were presented in red (the optimal color for this cell) at the cell's receptive field center (0.32° left of and 1.32° below fixation). (**b**) Gaussian shape-tuning function describing the response pattern in (**a**). The vertical axis represents boundary curvature, and the horizontal axis represents angular position of boundary fragments with respect to the shape's center of mass. The color scale (right) indicates normalized predicted response. The tuning peak corresponds to sharp convex curvature (1.0) near the top of the shape (84.6°). (**c**) Comparison of observed responses to responses predicted by the Gaussian tuning function, for the heart-shaped stimulus at eight orientations. Gray-level scale is the same as in (**a**). (**d**) Auxiliary test of object-centered position tuning for a different neuron.

rounding the stimulus icons. Darker backgrounds correspond to higher response rates (see scale bar). This cell responded to a variety of shapes with convex boundary curvature near the top. Sharper (higher, more acute) convex curvature was especially effective. This is most easily visualized by comparing responses to different orientations of the same shape across rows in **Fig. 1a**. In every case, orientations with sharper convex curvature at the top elicited stronger responses.

This response pattern was quantified with a 2D Gaussian tuning function (**Fig. 1b**). The horizontal axis represents angular position of the boundary fragment relative to the object's center of mass, with 0° corresponding to right, 90° corresponding to top, and so on. The vertical axis represents boundary curvature, with negative values corresponding to concavity and positive values corresponding to convexity. Curvature is defined as change in tangent angle per unit contour length (defined relative to stimulus size). Because curvature becomes infinite for tangent discontinuities (angles), we used a squashing function to map curvature to a range from –1.0 (sharp concave angles) to 1.0 (sharp convex angles). Boundary curvature in our stimulus set ranged from –0.3 to 1.0. Color (see scale bar) is used to represent tuning function amplitude—that is, predicted normalized response strength at each point on the curvature × angular position domain.

To derive the optimum tuning function, each stimulus was decomposed into four to eight contour fragments of relatively constant curvature. For each fragment, average curvature was measured, and angular position was calculated for the fragment center relative to the shape's center of mass. As an example, the crescent shape at the top left of **Fig. 1a** was decomposed into two sharp convexities (near the top and left), a broad convexity (lower right) and a concavity (upper left). During the iterative fitting procedure, the predicted response for each stimulus was based on whichever contour fragment fell closest to the Gaussian tuning function peak—whichever fragment predicted the strongest response. (Similar results were obtained when the predicted response was instead summed across fragments.) The tuning function was adjusted to minimize the sum of squared errors between observed and predicted responses. Further details of the fitting procedure have been described previously[20].

The tuning function peak (red) for this cell was at curvature 1.0 (sharp convex) and angular position 84.6° (slightly to the right of top center) (**Fig. 1b**). Responses predicted by the tuning function were strongly correlated ($r = 0.73$) with the observed responses shown in **Fig. 1a**. This is exemplified by the similarity between observed and predicted responses to different orientations of the heart-shaped stimulus (**Fig. 1c**). Both observed and predicted responses were highest for sharp convexities at the top and moderate for blunter convexities at the top or convexities near the upper right.

The importance of angular position as a tuning dimension was verified with an auxiliary test in which the position of the optimum boundary curvature relative to object center was parametrically varied (**Fig. 1d**). This example cell was tuned for sharp convexity near the lower right (315°). Tuning for angular position was maintained across different orientations of the sharp convexity (columns). Angular position tuning was also consistent across absolute position on the retina (data not shown). Similar results were obtained for most neurons tested in this way, as previously reported[20].

For each cell in our sample of 109 V4 neurons, we derived a curvature × position tuning function like that shown in **Fig. 1b**[20]. The tuning functions were highly predictive of observed responses, much more so than edge orientation tuning functions. In every case, we also derived a tuning function from a randomly selected subset of our stimuli, and found that this successfully

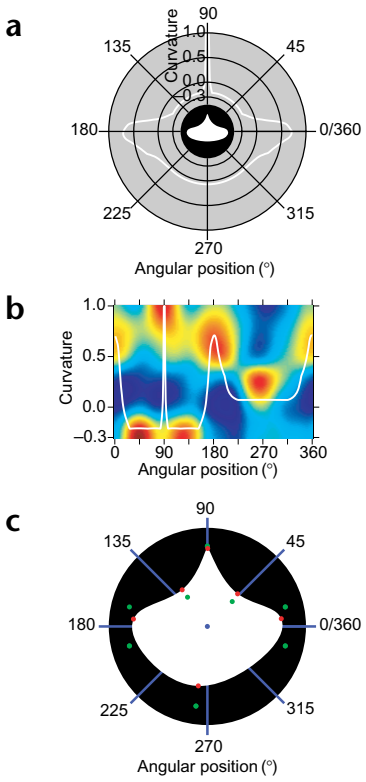

Fig. 2. Population response to an example shape. (**a**) The white line represents this shape's boundary curvature as a function of angular position, shown here in polar format centered around the shape's center of mass to highlight the correspondence with boundary features. (**b**) Estimated population response across the curvature × position domain (colored surface) with the veridical curvature function superimposed (white line). A Cartesian plot is used here because a polar plot would distort peak width in the population response. The surface was thresholded by subtracting the minimum value from all points, so amplitude varies from 0 (minimum) to 1 (maximum). (**c**) Reconstruction of the geometric shape from the population surface in (**b**).

predicted responses to the remaining stimuli, showing that the analysis generalizes to shapes not used in the fitting procedure[20]. Many neurons were additionally sensitive to more complex aspects of boundary conformation. For example, the cell in **Fig. 1** responded best when sharp convex curvature near the top was flanked by concave curvature on either side. This kind of response pattern can be more accurately predicted with tuning functions in a higher-dimensional domain[20]. As a first approach to population analysis, however, we have focused on the simpler curvature × angular position domain. In this domain, tuning function correlation values ranged from 0.11 to 0.81. All cells were included in the population analysis.

We used the curvature × position tuning functions (as in **Fig. 1b**) to estimate the V4 population response to each shape in our stimulus set. An example analysis is presented for the 'squashed raindrop' shape (**Fig. 2a**, center). The veridical curvature × position function for this stimulus is represented by the white line in the surrounding polar plot. This function has peaks and troughs corresponding to the major features of the shape: a medium convex peak at 0° (right), a concave trough at 45° (upper right), a sharp convex peak at 90° (top), and so forth. Our analysis was designed to determine whether similar curvature/position information was represented in the neural responses.

We scaled each cell's tuning peak by its response to the shape in question. Thus, each cell 'voted' for its preferred boundary fragment with a strength proportional to its response rate. The cell in **Fig. 1** responded strongly to the squashed raindrop shape, so the height of its scaled tuning peak was near 1.0. The entire set of scaled tuning peaks defined a surface representing coding strength for various combinations of curvature and position. We used a 2D Gaussian function to smooth that surface (standard deviation of 0.125 in the curvature dimension and 0.33 radians or 19° in the angular position dimension; the exact values were not critical to the outcome). We normalized (divided) this surface by an identical Gaussian convolution with the non-scaled (unit) point functions to help correct for uneven sampling across the curvature × position domain. The final result was an estimate of coding strength across the curvature × position domain for a hypothetical population of V4 neurons with identical receptive field centers responding to a single shape at a single position. In the actual experiments, of course, stimulus position was adjusted for each cell, over an eccentricity range of 0.0° to 6.6°. Given the variability in V4 receptive field position, it would be impractical to sample responses to a single stimulus position from a large enough number of cells. Thus, one assumption behind our analysis is that shape-coding mechanisms are similar across the tested eccentricity range.

The population surfaces derived in this manner contained peaks corresponding to the major boundary features in the stimulus shapes. However, there were still major amplitude differences between peaks due to sparse sampling in some regions of the curvature × position domain. There was no statistical correlation between curvature and angular position (linear-circular correlation coefficient[33] = 0.037, $P = 0.134$ by randomization test), and there is no *a priori* reason to expect that certain curvature values would be better represented at particular angular positions. Thus, the uneven distribution of tuning peaks was an artifact of our limited sample size.

To overcome this sampling limitation and produce a more accurate estimate of the true population response, we made the assumption, for each cell we studied, that with a large enough sample we would have encountered other cells with similar curvature tuning properties at different angular positions. We therefore replicated each cell's tuning function seven times, rotating across the angular position domain at 45° intervals. We assigned observed responses to each hypothetical cell by likewise rotating the original observed response rates across stimulus orientation. (Each stimulus was tested at 45° orientation intervals, so we could predict exactly how the hypothetical cell would respond.) Thus, our final population surfaces were based on an estimated sample of 872 cells, under the assumption that curvature representation is uniform across angular position. This approach produced response surfaces closer to what would have been obtained if a very large number of V4 neurons had been sampled. The results were qualitatively similar to those obtained without replication of tuning functions. The analyses presented here are based on the larger, estimated population.

Application of this analysis to the squashed raindrop shape yielded an estimated population surface (**Fig. 2b**) containing peaks corresponding to all the major boundary features: the sharp convexity at 90°, the medium convexities at 0° and 180°, the broad convexity at 270° and the concavities at 45° and 135°. The cell in **Fig. 1** contributed to the convex peak at 90° by virtue of its strong response to this shape. The veridical (stimulus-

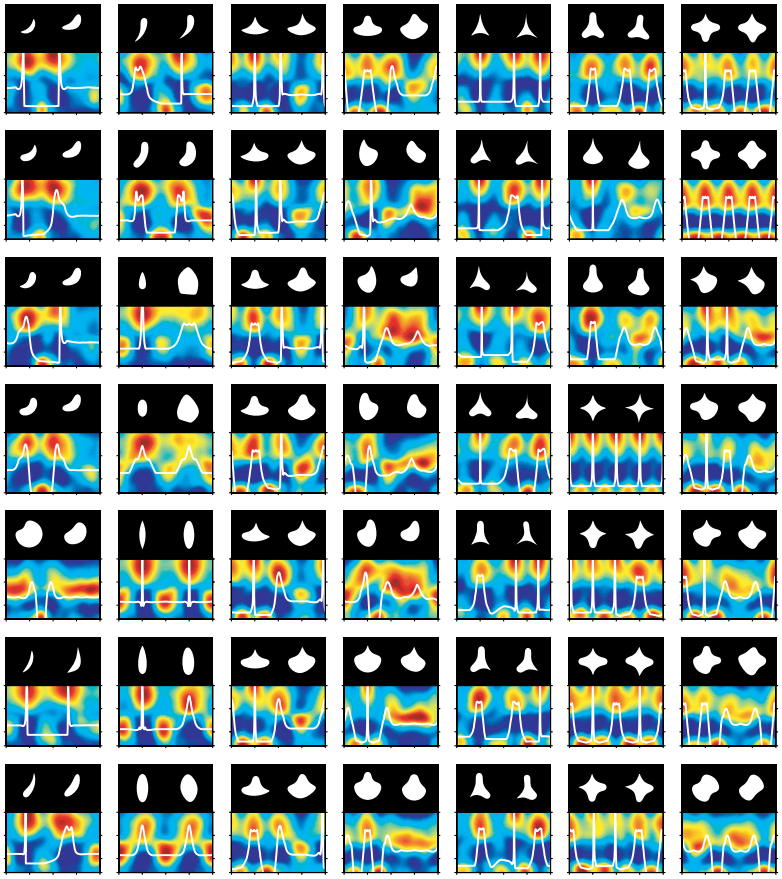

Fig. 3. Population responses to the 49 basic shapes, each at one orientation. The original shapes are shown as white icons against a black background at the top left of each panel. Reconstructed shapes are shown at the top right. Estimated population responses (colored surfaces) and veridical curvature functions (white lines) are shown at the bottom. Axes and scale for the population plots are the same as in **Fig. 2b**.

derived) shape function (white line) is replotted for comparison. This is a 1D circular function, with a single, exact curvature value at all angular positions. The summed population activity (colored surface) is a 2D function of curvature and angular position, in which amplitude corresponds to likelihood that a similar curvature/angle combination is present. As in other neural coding situations, graded peaks in the population activity approximate exact stimulus values in the environment. The novel aspect of our analysis is the multi-peaked representation of multiple stimulus components. The accuracy of this representation can be assessed by comparing the peaks in the population surface (red) to the veridical curvature values (white line) at the corresponding angular positions. The population peaks fall close to the extrema in the curvature function, and thus represent the salient boundary features in the shape. The most notable discrepancy is the peak near 270°, which has a curvature value of 0.25, somewhat higher (sharper) than the veridical curvature value of 0.075.

Another way to evaluate the population representation is to 'decode' it back into geometric space. We reconstructed the squashed raindrop shape (**Fig. 2c**) from the population surface by using B-spline[34] control points (green dots) to define a smooth (no tangent or curvature discontinuities) boundary with curvature extrema (red dots) matching the population surface peaks. The population peaks were calculated by sliding a window (45° angular position × 0.5 curvature) across the population surface to identify local peaks with values above 0.4. We used a single control point to approximate each peak with curvature less than 0.5, adjusting the radial position to produce a local curvature extremum matching the population curvature and angular position values. (Positions close to the center produced concavities, more eccentric positions produced convexities.) Two control points were necessary to define sharper curvatures greater than 0.5 (the two control points at the top of **Fig. 2c** overlap because of the high curvature value). The radial position of these control point pairs was fixed at the maximum distance from the center (making the reasonable assumption that sharper convexities would be furthest from the center). Their angular separation was varied to produce a local convexity matching the population curvature value. Radial and angular control point positions were iteratively adjusted to minimize differences between the reconstructed shape's curvature extrema and the population peaks. This reconstruction method is not meant to have any biological relevance; it is only an expedient for producing a shape with boundary features approximating the peaks in the response surface, in order to visualize the accuracy of the population representation.

In the reconstruction of the squashed raindrop shape (**Fig. 2c**), all the major elements of the original stimulus are visible. This shows the effectiveness of the population representation in curvature × angular position space, at least for this type of moderately complex shape. The most obvious discrepancy in **Fig. 2c** is the sharper curvature near the bottom, as expected due to the high value of the population peak at 270°. This error may reflect a sampling problem in our data set, as we found relatively few cells with tuning peaks in this curvature range[20]. In addition, minor offsets in the angular positions of the peaks skew the symmetry of the reconstructed shape.

We generated population response estimates and reconstructions for all 49 shapes in our stimulus set (at one orientation each; **Fig. 3**). Each panel shows the original stimulus at the top left, the reconstructed stimulus at the top right, and the population estimate (colored surface) at the bottom, with the veridical curvature × position function (white line) superimposed. In most cases, all the salient boundary features are represented in the population surface and reproduced in the reconstruction, albeit with some degree of error in curvature and angular position. Specifically, distinctions between sharp and medium convexities are not always represented, and the curvature of broad convexities tends to be overestimated. In some cases, adjacent peak values were not completely geometrically compatible, so that curvatures and angular positions in the reconstructions could not perfectly match those in the population surface. These discrepancies may be partially due to sampling bias and response measurement error, and they may to some extent reflect poor stimulus perception by the animals. In addition, however, the errors suggest that our 2D domain

does not perfectly capture the information in V4 population responses. It seems likely that curvature and angular position are related to inherent V4 tuning dimensions, but any number of similar shape parameterizations might be more accurate. More than two dimensions would obviously be required to explain all the variance in V4 shape responses, and to describe more complex objects, 3-D objects, objects with internal structure or objects defined by color and texture properties.

We further assessed the accuracy of the population signals by computing (for each shape) the mean absolute difference between the population peaks and the nearest points on the veridical curvature × position function (**Fig. 4**). Mean absolute differences were small in both the curvature dimension (vertical axis; median 0.0704) and the angular position dimension (horizontal axis; median 4.04°). The **Fig. 2** example stimulus is plotted in red.

Discussion

Previous studies of population responses in the ventral pathway have focused on general characteristics of population-level information. Some groups have used information theory to measure the representational capacity of neural populations in temporal visual cortex[35,36]. For instance, it has been shown that population responses in anterior inferotemporal cortex (IT) to photographic face stimuli are significantly related to physical similarity based on distance measurements between face features[37]. This reflects 'second-order isomorphism'[38]—that similar shapes evoke similar neural responses. IT population responses to parametrically varying novel shapes also show second-order isomorphism[39]. It has also been shown that shape discrimination training leads to greater distances between learned stimuli in IT population response space[40]. In monkeys trained to perform categorical shape discrimination, the IT population signal carries more information about parts-level features that distinguish the learned categories[18]. This implies a parts-based representation of the type described here. Other studies of IT shape responses also suggest a parts-based coding scheme[14–17,21].

These studies elucidated general characteristics of population responses, but here we sought to describe the actual population code in area V4—which neurons represent what elements of shape information and in what way. Our results suggest that V4 neurons encode shapes at least partially in terms of their constituent boundary features, through graded tuning for conformation and object-centered position[32]. Information about a complete shape is available from the multiple peaks in the population activity profile.

Our analysis extrapolates from single-peak methods such as vector summation, devised by K.O. Johnson[41] and first implemented by Georgopoulos and colleagues[27]. The vector-sum method is mathematically equivalent to finding the population activity peak by fitting a cosine function[30]. Other methods for extracting a single estimated value from a neural population profile have also been explored[29–31,42,43]. Our approach here represents a natural extension of these methods: extracting multiple values from a complex population pattern representing a multicomponent stimulus. This type of multiplexed population signal is bound to be important in representing other types of complex stimuli. For example, a recent study demonstrated that the visual system depends on the shape of the MT population peak to represent multiple overlapping motion directions[44].

Our results imply a 'structural' coding scheme, consistent with the notion of representation by parts or components. In models based on this idea, shapes are described in terms of the conformations and relative positions (and/or connectivity) of their simpler components[9–13]. A recent model demonstrates that parts-level coding could rely on retinotopic (rather than relative or object-centered) position[45]. Our previous data imply relative position coding for parts on a local level, within the V4 receptive field[20]. On a larger scale, however, V4 neurons convey retinotopic position information.

For convenience, structural models are sometimes discussed in terms of stereotyped parts or 'primitives', but our results imply continuous part representation by neurons with graded tuning in quantifiable shape dimensions (a 'multidimensional feature space'[38]). The advantages of parts-based representation are its robustness to variations in the retinal image of an object and its alphabet-like power to encode an infinite variety of shapes. Most parts-based theories envision a hierarchical progression of parts complexity through a sequence of processing stages. Boundary fragments constitute an appropriate level of parts complexity for an intermediate processing stage like V4. More complex parts are encoded in IT, the next stage in the ventral pathway[14–16,46,47].

The alternative to parts-based or structural representation of shape is holistic representation, where each neuron carries information about an entire object or scene rather than just one part[21,38]. Intermediate coding schemes are also possible, where each neuron carries information about the entire object but more about some parts than others. If shape representation in V4 were mainly holistic, our population analysis would have yielded flat or noisy surfaces rather than multiple peaks corresponding to separate parts. The large amounts of response variance not explained by the Gaussian part-tuning functions, ranging from 34% ($r = 0.81$) to 99% ($r = 0.11$), could have reflected holistic coding—different responses to every combination of parts and thus to every shape. However, the fact that strong peaks consistently emerged in the weighted population sums argues for parts-based shape representation in V4.

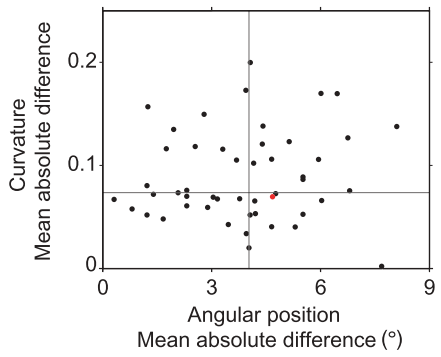

Fig. 4. Population coding accuracy. Each of the 49 basic shapes (shown in **Fig. 3** is represented as a point. The red point corresponds to the squashed raindrop shape in **Fig. 2**. The vertical axis represents mean absolute difference between population surface peaks and the veridical boundary function in the curvature dimension. (Our squashed curvature values have no units, but are derived from absolute curvature, which is defined as change in tangent angle per unit contour length.) The horizontal axis represents mean absolute difference in the angular position dimension. These averages were taken across 3–8 peaks, depending on the population surface (see **Fig. 3**). For this analysis, population surface peaks were identified as local maxima within a 90° (angular position) × 0.25 (curvature) sliding window. Median differences are represented by thin lines.

It should be noted that our shape reconstruction method (**Figs. 2c** and **3**) is just a visualization tool and not a hypothesis about how the visual system uses population information. The visual system has no need to decode its representations back into geometric space. On the contrary, encoding into dimensions like curvature and relative position presumably advances the transformation towards a more invariant and robust object representation in IT cortex. Also, our simple reconstruction method is not an optimal translation of the population surface, which contains more information than just discrete peak positions[44]. For example, some population peaks were very broad in the angular position dimensions, representing long expanses of broad convex curvature. In our analysis this produced multiple discrete peaks at 90° intervals, but the visual system must use such information in a more continuous fashion.

Our study has a number of limitations. As in any experiment on shape, the stimulus set sampled only a small portion of the potential space. Our stimuli were silhouettes with no internal structure, so our results address coding of 2D shape boundaries only. Boundary curvature was sampled at only a few levels, so we may have failed to pinpoint the true tuning peak for many cells. In particular, we did not have straight edges (0 curvature) in our stimuli, so our conclusions are limited to shapes composed of curved boundary fragments. The curvature fragments in our shapes were mostly oriented with respect to center of mass—all the convex projections radiated from a single center. Shapes containing curves at other orientations would require a higher-dimensional description. Thus, our conclusions are limited to a particular class of shapes. The general method, however, could be extended to other shape characteristics and other object categories.

Methods

Data collection. We recorded the activity of single V4 neurons in two female rhesus monkeys (*Macaca mulatta*). During each experimental session, the monkey was seated in front of a computer monitor with the head immobilized by means of a custom-built titanium head post. The animal was trained to fixate a 0.1° white spot within 0.5° of visual angle for a period of 3.75 s. Eye position was monitored using the scleral search coil method[48]. While the animal performed the fixation task, epoxy-coated tungsten electrodes (A-M Systems, Carlsburg, Washington) were used to record neural activity from individual V4 neurons in the lower parafoveal representation on the prelunate gyrus and adjoining sulcal banks. We isolated a total of 409 V4 neurons. We chose 222 for further study because preliminary tests showed that they belonged to the subpopulation of V4 cells sensitive to complex shape[24]. Other cells were tuned only for bar/edge orientation or were not responsive to any of our stimuli. Of the 222 complex shape cells, we were able to study 109 long enough to present at least three (usually five) repetitions of each stimulus. All animal procedures conformed to National Institutes of Health and USDA guidelines and were approved by the Johns Hopkins University Animal Care and Use Committee. Detailed methodology has been described previously[20].

Stimuli. The stimulus set (**Fig. 1a**) comprised 49 complex shapes constructed by systematically combining convex and concave boundary fragments. Each stimulus was presented at 2, 4 or 8 orientations depending on rotational symmetry. Stimuli were rendered in the cell's optimal color against a gray background and presented at the center of the cell's receptive field. Stimulus size was scaled with average receptive field size at the cell's eccentricity[49], so as to cover approximately three quarters of the receptive field diameter. (Scaling maintains visibility at greater eccentricities.) During each behavioral trial, a sequence of five randomly chosen stimuli were flashed for periods of 500 ms each, separated by 250-ms interstimulus intervals. For most cells, each stimulus was presented five times. Details of stimulus construction and presentation are described elsewhere[20].

Acknowledgments

Technical assistance was provided by W. Nash and B. Sorenson. A.J. Bastian, K.O. Johnson, T. Poggio and M. Riesenhuber made helpful comments on previous versions of the manuscript. This work was supported by the National Eye Institute and by the Pew Scholars Program in the Biomedical Sciences.

Competing interests statement

The authors declare that they have no competing financial interests.

RECEIVED 28 AUGUST; ACCEPTED 15 OCTOBER 2002

1. Ungerleider, L. G. & Mishkin, M. *Analysis of Visual Behavior* (eds. Ingle, D. G., Goodale, M. A. & Mansfield, R. J. Q.) 549–586 (MIT Press, Cambridge, Massachusetts, 1982).
2. Felleman, D. J. & Van Essen, D. C. Distributed hierarchical processing in the primate cerebral cortex. *Cereb. Cortex* **1**, 1–47 (1991).
3. Milner, P. M. A model for visual shape recognition. *Psychol. Rev.* **81**, 521–535 (1974).
4. Selfridge, O. G. *The Mechanization of Thought Processes* (H. M. Stationary Office, London, 1959).
5. Sutherland, N. S. Outlines of a theory of visual pattern recognition in animals and man. *Proc. R. Soc. Lond. B Biol. Sci.* **171**, 297–317 (1968).
6. Barlow, H. B. Single units and sensation: a neuron doctrine for perceptual psychology? *Perception* **1**, 371–394 (1972).
7. Hubel, D. H. & Wiesel, T. N. Receptive fields of single neurones in the cat's striate cortex. *J. Physiol. (Lond.)* **148**, 574–591 (1959).
8. Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J. Physiol. (Lond.)* **195**, 215–243 (1968).
9. Marr, D. & Nishihara, H. K. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc. R. Soc. Lond. B Biol. Sci.* **200**, 269–294 (1978).
10. Hoffman, D. D. & Richards, W. A. Parts of recognition. *Cognition* **18**, 65–96 (1984).
11. Biederman, I. Recognition-by-components: a theory of human image understanding. *Psychol. Rev.* **94**, 115–147 (1987).
12. Dickinson, S. J., Pentland, A. P. & Rosenfeld, A. From volumes to views: an approach to 3-D object recognition. *CVGIP: Image Understanding* **55**, 130–154 (1992).
13. Riesenhuber, M. & Poggio, T. Hierarchical models of object recognition in cortex. *Nat. Neurosci.* **2**, 1019–1025 (1999).
14. Tanaka, K., Saito, H., Fukada, Y. & Moriya, M. Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *J. Neurophysiol.* **66**, 170–189 (1991).
15. Fujita, I., Tanaka, K., Ito, M. & Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**, 343–346 (1992).
16. Tsunoda, K., Yamane, Y., Nishizaki, M. & Tanifuji, M. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nat. Neurosci.* **4**, 832–838 (2001).
17. Wang, Y., Fujita, I. & Murayama, Y. Neuronal mechanisms of selectivity for object features revealed by blocking inhibition in inferotemporal cortex. *Nat. Neurosci.* **3**, 807–813 (2000).
18. Sigala, N. & Logothetis, N. K. Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature* **415**, 318–320 (2002).
19. Pasupathy, A. & Connor, C. E. Responses to contour features in macaque area V4. *J. Neurophysiol.* **82**, 2490–2502 (1999).
20. Pasupathy, A. & Connor, C. E. Shape representation in area V4: position-specific tuning for boundary conformation. *J. Neurophysiol.* **86**, 2505–2519 (2001).
21. Baker, C. I., Behrmann, M. & Olson, C. R. Impact of learning on representation of parts and wholes in monkey inferotemporal cortex. *Nat. Neurosci.* **5**, 1210–1216 (2002).
22. Desimone, R. & Schein, S. J. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *J. Neurophysiol.* **57**, 835–868 (1987).
23. Gallant, J. L., Braun, J. & Van Essen, D. C. Selectivity for polar, hyperbolic and Cartesian gratings in macaque visual cortex. *Science* **259**, 100–103 (1993).
24. Kobatake, E. & Tanaka, K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J. Neurophysiol.* **71**, 856–867 (1994).
25. Zeki, S. Color coding in rhesus monkey prestriate cortex. *Brain Res.* **53**, 422–427 (1973).
26. Hanazawa, A. & Komatsu, H. Influence of the direction of elemental luminance gradients on the responses of V4 cells to textured surfaces. *J. Neurosci.* **21**, 4490–4497 (2001).
27. Georgopoulos, A. P., Caminiti, R., Kalaska, J. F. & Massey, J. T. Spatial coding of movement: a hypothesis concerning the coding of movement direction by motor cortical populations. *Exp. Brain Res.* **7** (Suppl.), 327–336 (1983).
28. Lewis, J. E. & Kristan, W. B. Jr. A neuronal network for computing population vectors in the leech. *Nature* **391**, 76–79 (1998).
29. Salinas, E. & Abbott, L. F. Vector reconstruction from firing rates. *J. Comput. Neurosci.* **1**, 89–107 (1994).

30. Deneve, S., Latham, P. E. & Pouget, A. Reading population codes: a neural implementation of ideal observers. *Nat. Neurosci.* **2**, 740–745 (1999).
31. Zhang, K., Ginzburg, I., McNaughton, B. L. & Sejnowski, T. J. Interpreting neuronal population activity by reconstruction: unified framework with application to hippocampal place cells. *J. Neurophysiol.* **79**, 1017–1044 (1998).
32. Olson, C. R. & Gettner, S. N. Brain representation of object-centered space. *Curr. Opin. Neurobiol.* **6**, 165–170 (1996).
33. Mardia, K. V. Linear-circular correlation coefficients and rhythmometry. *Biometrika* **63**, 403–405 (1976).
34. Bartels, R. H., Beatty, J. C. & Barsky, B. A. *An Introduction to Splines for Use in Computer Graphics and Geometric Modeling* (Morgan Kaufmann, Los Altos, California, 1987).
35. Rolls, E. T., Treves, A. & Tovee, M. J. The representational capacity of the distributed encoding of information provided by populations of neurons in primate temporal visual cortex. *Exp. Brain Res.* **114**, 149–162 (1997).
36. Gochin, P. M., Colombo, M., Dorfman, G. A., Gerstein, G. L. & Gross, C. G. Neural ensemble coding in inferior temporal cortex. *J. Neurophysiol.* **71**, 2325–2337 (1994).
37. Young, M. P. & Yamane, S. Sparse population coding of faces in the inferotemporal cortex. *Science* **256**, 1327–1331 (1992).
38. Edelman, S. *Representation and Recognition in Vision* (MIT Press, Cambridge, Massachusetts, 1999).
39. Op de Beeck, H., Wagemans, J. & Vogels, R. Inferotemporal neurons represent low-dimensional configurations of parameterized shapes. *Nat. Neurosci.* **4**, 1244–1252 (2001).
40. Kobatake, E., Wang, G. & Tanaka, K. Effects of shape-discrimination training on the selectivity of inferotemporal cells in adult monkeys. *J. Neurophysiol.* **80**, 324–330 (1998).
41. Mountcastle, V. B. The parietal system and some higher brain functions. *Cereb. Cortex* **5**, 377–390 (1995).
42. Seung, H. S. & Sompolinsky, H. Simple models for reading neuronal population codes. *Proc. Natl. Acad. Sci. USA* **90**, 10749–10753 (1993).
43. Zemel, R. S., Dayan, P. & Pouget, A. Probabilistic interpretation of population codes. *Neural Comput.* **10**, 403–430 (1998).
44. Treue, S., Hol, K. & Rauber, H. J. Seeing multiple directions of motion-physiology and psychophysics. *Nat. Neurosci.* **3**, 270–276 (2000).
45. Edelman, S. & Intrator, N. (Coarse coding of shape fragments) + (retinotopy) approximately = representation of structure. *Spat. Vis.* **13**, 255–264 (2000).
46. Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *J. Neurosci.* **4**, 2051–2062 (1984).
47. Janssen, P., Vogels, R. & Orban, G. A. Macaque inferior temporal neurons are selective for disparity-defined three-dimensional shapes. *Proc. Natl. Acad. Sci. USA* **96**, 8217–8222 (1999).
48. Robinson, D. A. A method of measuring eye movement using a scleral search coil in a magnetic field. *IEEE Trans. Bio-Med. Electron.* **10**, 137–145 (1963).
49. Gattass, R., Sousa, A. P. & Gross, C. G. Visuotopic organization and extent of V3 and V4 of the macaque. *J. Neurosci.* **8**, 1831–1845 (1988).