Neural clusters encoding curvature and corners emerged in macaque V4

Rundong Jiang[1,2,3] Ming Li[4] & Shiming Tang[1,2,3*]

ABSTRACT

As an intermediate stage of the ventral visual pathway, V4 plays a crucial role in transforming basic orientation information into higher-ordered object representations. However, the neural mechanisms underlying the encoding of simple, complex and intermediate features in V4 remain poorly understood. Using two-photon calcium imaging in awake macaques, we recorded the responses of large populations of V4 neurons to thousands of natural images. To understand these high-dimensional data sets, we performed dimensional reduction analysis on the neuronal population responses. We found orthogonal dimensions encoding orientation and more complex features, with curves and corners represented separately. These distinct feature dimensions were encoded by spatially clustered subpopulations of neurons organized in iso-orientation and curvature domains. Using synthetic stimuli, we confirmed that curvature and corner selective neurons in V4 were tuned to integral features rather than combinations of local orientation compartments. Our results show that curves and corners are encoded by neurons clustered into functional domains separate from orientation. This functionally specific population architecture may serve to facilitate local computations that expedite more complex shape and object representations at higher levels of the ventral pathway.

[1] Peking University School of Life Sciences, Beijing, China. [2] Peking-Tsinghua Center for Life Sciences, Beijing, China. [3] IDG/McGovern Institute for Brain Research at Peking University, Beijing, China. [4] Beijing Normal University, Beijing, China. Correspondence should be addressed to S. T. (tangshm@pku.edu.cn).

INTRODUCTION

The visual system faces the daunting task of combining ambiguous local patterns of contrast into robust and spatially extensive complex shape and object representations. Such shape information is predominantly processed along the ventral pathway (V1, V2, V4, IT)[1,2]. At early stages of this cortical pathway, neurons are tuned to local orientation[3,4], or in some cases combinations of orientations[5,6]. Orientation is functionally organized into iso-orientation domains that form pinwheel structures in V1[7]. At later stages like IT, neurons are selective for complex objects or object categories[8-11]. These complex objects are functionally organized within IT by combinations of structurally separated feature columns[9,12,13]. This functional transformation from local orientation to complex object representation is dependent on an intermediate processing stage, cortical area V4[14].

The functional organization of V4 has previously been visualized by intrinsic signal optical imaging, and cortical representation of orientation, colour and spatial frequency has been systematically studied and demonstrated[15,16], suggesting that V4 bears some similarities with V1 in representing lower-order visual information. However, how complex feature selective neurons in V4 are spatially organised and whether feature columns found in IT also exist in V4 remain unclear. Because intrinsic imaging is spatially and temporally limited, it cannot measure selective responses of single neurons. Using electrophysiology, early studies in V4 using bar and grating stimuli found that V4 neurons are tuned for orientation, size and spatial frequency[17]. Subsequent studies revealed V4 selectivity for complex gratings and shapes in natural scenes[18-20]. Systematic tuning of V4 neurons for shape segments such as curves and angles as well as combination of these segments was discovered using parametric stimulus sets consisting of complex features[21-26]. Image statistics have been used to reconstruct and simulate the receptive field of neurons from their responses to natural textures[27,28]. More recently, temporally varying heterogeneous fine-scale tuning within the spatial-temporal receptive field was reported[29,30]. Part of the continuing questions regarding V4 selectivity arises from the difficulties in relating studies that use either natural or synthetic images. Natural images minimise stimulus biases with their high dimensionality[31-33], whereas synthetic images limit their dimensionality, but by necessity bias any understanding to these dimensions alone. One fruitful avenue is to use natural image preferences to more carefully target further analysis using synthetic images[19,34,35]. Feature extraction and simplification from the preferred natural images helped to reveal neuronal selectivity for complex features in V1, V4 and IT.

In this study, we wished to address two questions that remain unresolved, what are the image features that identify V4 responses, and what are their functional and spatial relationships? To address this question, we performed two-photon calcium imaging in awake macaque cortex. This technique increases the spatial resolution for functional characterization at the single-cell level while enabling the

visualization of spatial distribution and clustering within the cortical population[36-38]. We recorded the responses of hundreds of V4 neurons to thousands of natural images. We used image statistics and dimensional reduction to better understand and interpret our data, finding that feature dimensions representing orientation and complex features like curves and corners are spatially clustered. We confirmed these findings using synthetic stimuli like gratings and contour features, and further investigated the detailed neuronal coding for more complex patterns. Our results provide evidence that neurons in upper layers of V4 are clustered into functional domains that relate lower and intermediate visual features during natural vision.

RESULTS

We injected AAV1-hSyn-GCaMP into V4 of two rhesus macaques — GCaMP5G for monkey A and GCaMP6s for monkey B. The cranial imaging window and head posts were implanted 1-2 months after viral injection[37]. During experiments, monkeys were required to perform a simple fixation task for 2.5s with head restrained. Stimuli were presented on an LCD monitor 45 cm away from monkeys' eyes (30 pixels/°) for 1 sec after 1.5 sec blank. Neuronal responses were recorded using two-photon calcium imaging, measuring differential fluorescence $\Delta F = F - F0$ (where F0 is the average fluorescence of 0-0.5 sec before stimulus onset, and F is the average of 0.5-1.25 sec after stimulus onset).

For monkey A, we randomly chose a recording site in V4 for two-photon imaging (**Fig. 1a-b**). The retinal eccentricity was 0.85° when mapped using drifting gratings. We recorded V4 layer II neurons' responses to 4136 natural images stimuli, each repeated at least 3 times. The natural images were downloaded from ImageNet and cropped into a 3° diameter round patch. We identified 291 active neurons responding selectively to the stimuli (**Supplementary Fig. 1**). We found a negative correlation between neuronal pairwise response correlation and spatial distance (**Fig. 1c**; Pearson correlation = -0.46), indicating that functionally similar neurons are spatially clustered. Furthermore, we found correlations between the population responses to some classes of stimuli, as indicated by the similarity of differential fluorescence images (**Fig. 1d-e**). This could be used to categorize the critical common stimulus features that certain subpopulations of neurons were selective for.

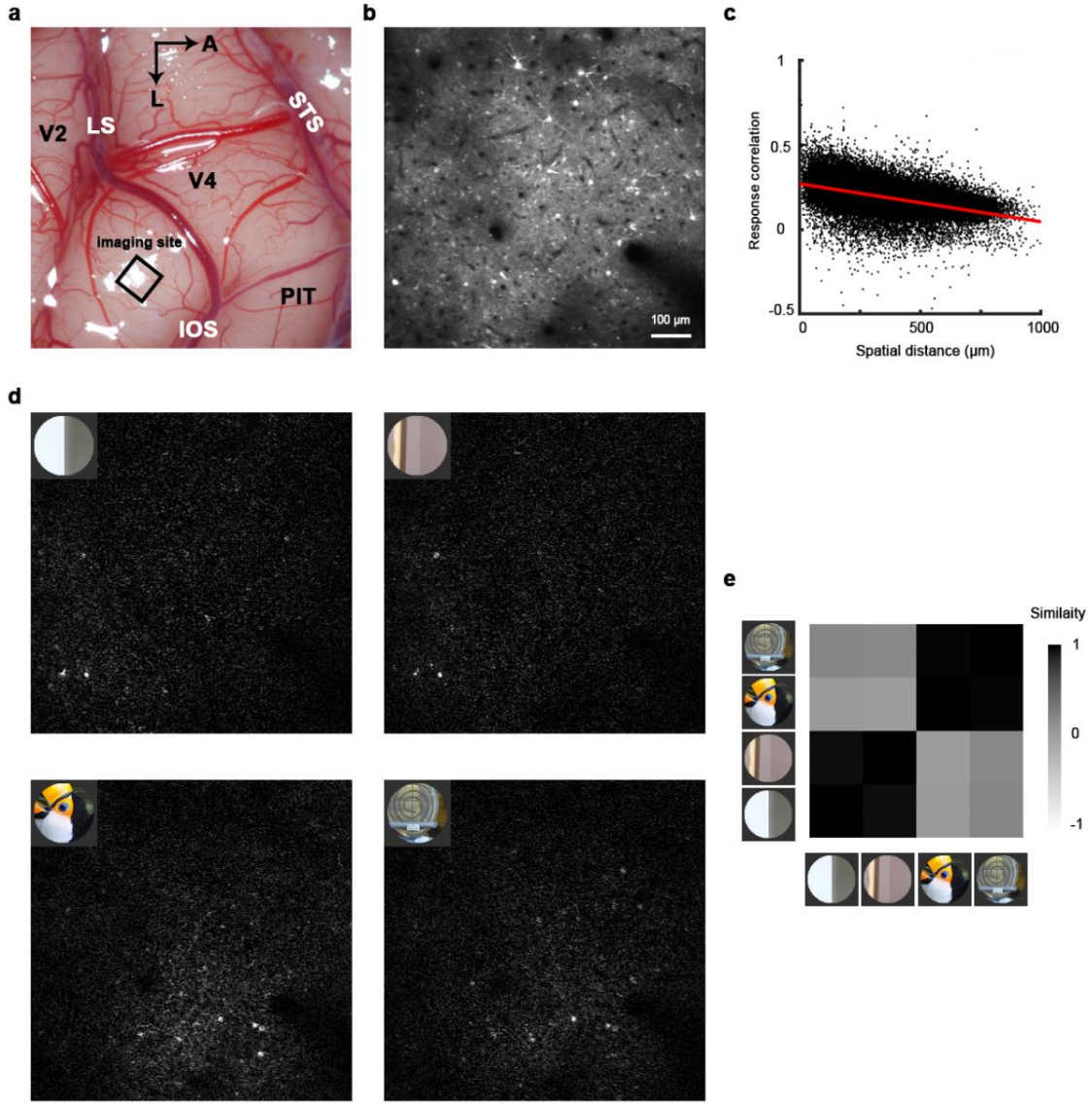

Figure 1 Responses of V4 layer 2 neurons to natural images stimuli. (**a**) The vascular map of monkey A. The site chosen for two-photon imaging is identified by a black box. A, anterior; L, lateral; LS, lunate sulcus; IOS, inferior occipital sulcus; STS, superior temporal sulcus. (**b**) Two-photon image of V4 layer 2 neurons expressing GCaMP5G (16x objective). (**c**) Scatterplot of neurons' pairwise response correlation against spatial distance, Pearson correlation = -0.46. (**d**) Population responses (differential fluorescence images, ΔF) evoked by natural image stimuli. (**e**) Some stimuli evoked very similar population responses. Similarity here was defined as the pairwise 2D-correlation of differential fluorescence images after applying a Gaussian filter.

MDS analysis demonstrates representation of orientation and complex features

To statistically characterize neuronal selectivity, it is necessary to find a parametric description of natural image stimuli. However, intuitive parameterization is difficult due to the high dimensionality of natural images. Dimensional reduction of stimulus space can be performed either according to the geometry of

the stimuli[39,40], or in a purely data-driven way based on neuronal responses[11,41]. Here we used the latter by performing multi-dimensional scaling (MDS) according to population responses evoked by 4136 natural image stimuli. Since we have already demonstrated the functional compartmentalization of neurons (**Fig. 1c**), as well as that some stimuli could evoke very similar differential fluorescence images (**Fig. 1d-e**), we constructed the distance matrix used for MDS by the pairwise 2D correlation of Gaussian-filtered differential fluorescence images as $D_{ij} = 1 - \mathrm{corr2}(\Delta F_i, \Delta F_j)$, so that the distances ranged from 0 to 2 (**Supplementary Fig. 2a**). The 4136 stimuli were then mapped to a 20 dimensional feature space by classical MDS with stress < 5%, and the polarity of each axis was re-defined (**Supplementary Fig. 2b-e**).

Because direct visualization of all 4136 stimuli by a scatter plot was difficult, we first examined the stimuli that had the largest coordinates on each feature dimensions (**Fig. 2a** and **Supplementary Fig. 2f**). Among these stimulus images empirically selected based on their coordinates, we noted some common geometric features were often shared. For dimension 1, many of these images contained a vertical orientation (**Fig. 2b**), and for dimension 2 and 4, horizontal and 45° orientations (**Fig. 2c-d**). This observation was consistent with our expectation that orientation, given its importance, could be statistically identified as a principal feature dimension. More interestingly, many of the images that had the largest coordinates in the forward direction of dimension 3 had circles or curves (**Fig. 2e**), while images with the largest coordinates in the reverse direction often had corners or polygonal lines (**Fig. 2f**). This appears to suggest an extra curvature dimension independent of orientation, in which smoothness distinguishes between curves and corners.

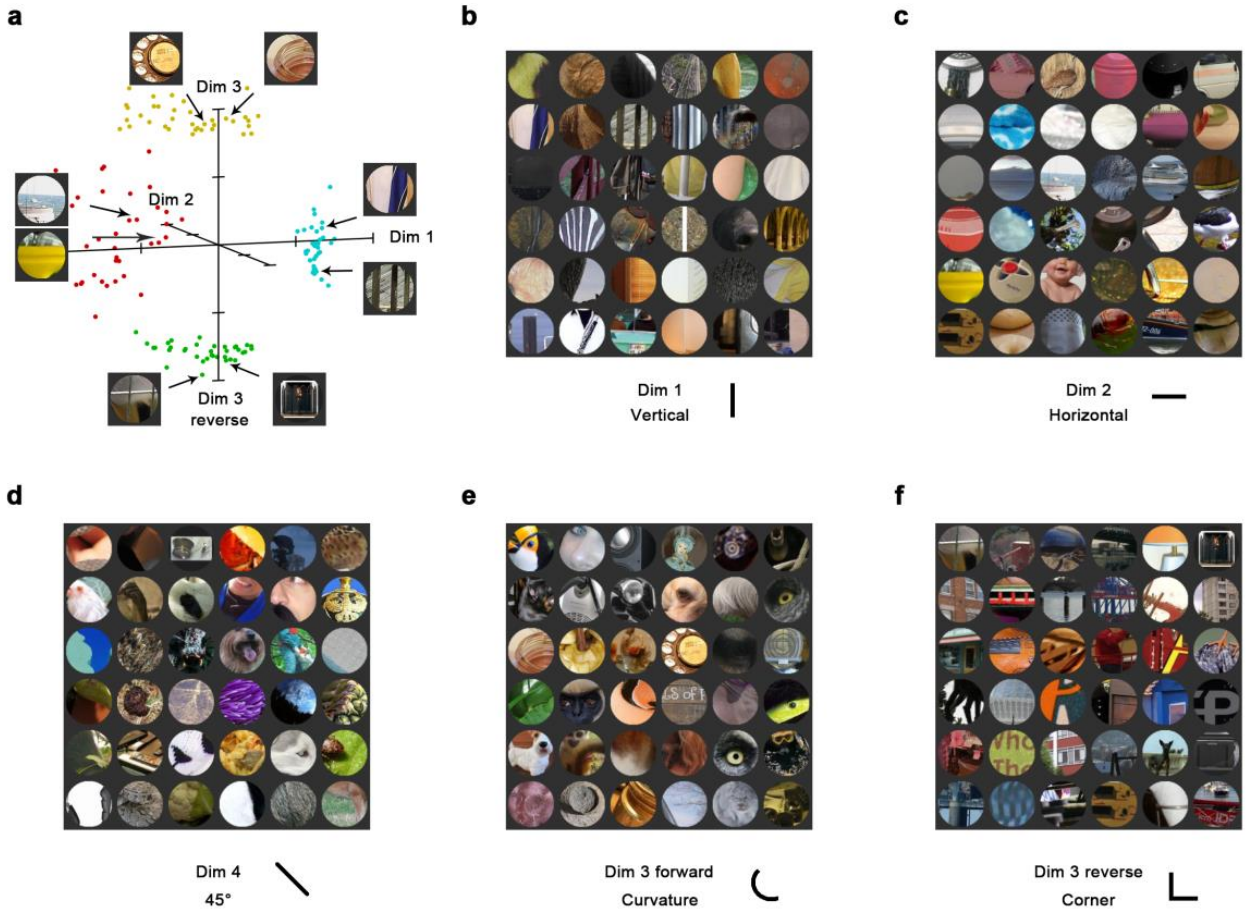

Figure 2 MDS analysis of neuronal responses to 4136 natural images stimuli. The distance matrix was calculated as 1 – corr2D, which is the pairwise 2D correlation of the ΔF images, and classical MDS was performed to map the 4136 stimuli into a 20-dimensional feature space. (**a**) Scatterplot of natural image stimuli in a feature space composed of the first three dimensions. Stimuli having the largest coordinates on dimension 1-3 and the reverse direction of dimension 3 were mapped. Two example stimuli were presented for each group. (**b**) The stimuli that had the largest coordinates on dimension 1; many of these images contained vertical orientation. (**c**) The stimuli that had the largest coordinates on dimension 2; many of these images contained horizontal orientation. (**d**) The stimuli that had the largest coordinates on dimension 4; many of these images contained 45° orientation. (**e**) The stimuli that had the largest coordinates on dimension 3; many of these images contained curves. (**f**) The stimuli that had the largest coordinates on the reverse direction of dimension 3; many of these images contained corners.

The disparate representation of curves and corners and its orthogonality to orientation implies that these features might possibly be encoded by different subpopulations of V4 neurons. Supporting this supposition, when we examined the preferred natural images of individual neurons, we found neurons selective for stimuli containing a particular orientation (Cell 38, **Fig. 3a**), or stimuli containing curves (Cell 140, **Fig. 3b**) or corners (Cell 182, **Fig. 3c**). The observed neuronal selectivity was consistent with its

tuning for the feature dimensions. For example, Cell 140 showed stronger responses to stimuli with large positive coordinates on dimension 3, the curvature dimension (**Fig. 3d**). We used response triggered average (RTA; averaging the dimensional coordinates weighted by the responses)[42] to characterize neuronal tuning along the feature dimensions. A positive RTA represents neurons with positive dimensional preferences. So, for example, cell 38 and cell 140 exhibited large positive RTA values for dimension 1 and 3 respectively, and Cell 182 large negative RTA values for dimension 3 (**Fig. 3e**). We also found that for each of the first four dimensions, neurons with high RTA were spatially clustered, and organized in different areas within the imaging site (**Fig. 3f** and **Supplementary Fig. 3**). This indicates that these feature dimensions representing simple and complex shape information are encoded by neuronal clusters in separate functional domains of V4.

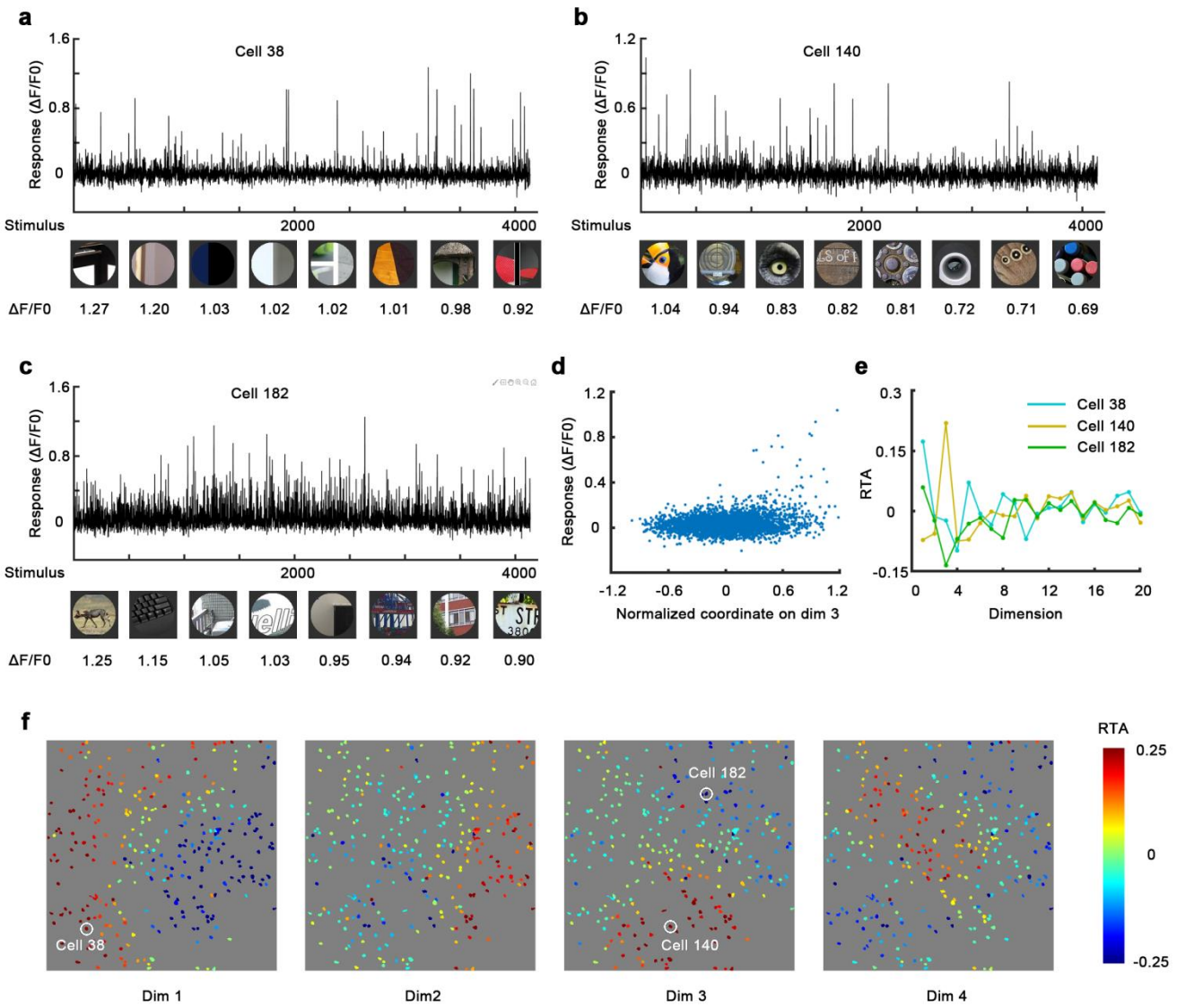

Figure 3 Neural tuning for feature dimensions. (**a**) responses of cell 38 to 4136 natural image stimuli. The preferred stimuli were displayed from left to right in descending rank order; many of its preferred stimuli contain vertical orientations. (**b** responses of cell 140; many of its preferred stimuli contain

curvature. (**c**) responses of cell 182; many of its preferred stimuli contain corners. (**d**) Scatter plot of cell 140's responses to natural image stimuli against their coordinates on dimension 3. Strong responses were evoked by stimuli whose coordinates were larger on dimension 3. (**e**) RTA (response triggered average, see Online Methods and ref.) was calculated for each neuron to characterize its tuning to the feature dimensions. Positive RTA means the neuron prefer stimuli with large coordinates on this dimension. Cell 38 and Cell 140 had large RTA on dimension 1 and 3, respectively. Cell 182 had large negative RTA on dimension 3. (**f**) Cell map of neuronal RTA on the first 4 feature dimensions. Neurons with large RTA on each feature dimension were spatially clustered.

Orientation representation derived by Fourier transformation and its consistency with MDS results

Although the previous results were indiciative of stimulus features (i.e. an oriented line) and the dominant MDS dimension (**Fig. 4a**), these features were not measured quantitatively. And so to better confirm these results, we first quantified the orientation information present in the natural image themselves. We performed a Fourier transform to extract the orientation spectrum of each stimulus to characterize its orientation components, and thus calculated neurons' orientation tuning (**Supplementary Fig. 4**). We then made the cell map of preferred orientation, which characterized the orderly organization of iso-orientation domains (**Fig. 4b**), and we found it was highly consistent with our MDS result. Vertical orientation corresponded with dimension 1, horizontal with dimension 2, and 45° with dimension 4. We also measured neuronal responses to orientated bars at the same recording site and found consistent structure (**Fig. 4c-d**), which appeared to be part of an orientation pinwheel when imaged at a larger scale (**Fig. 4c**).

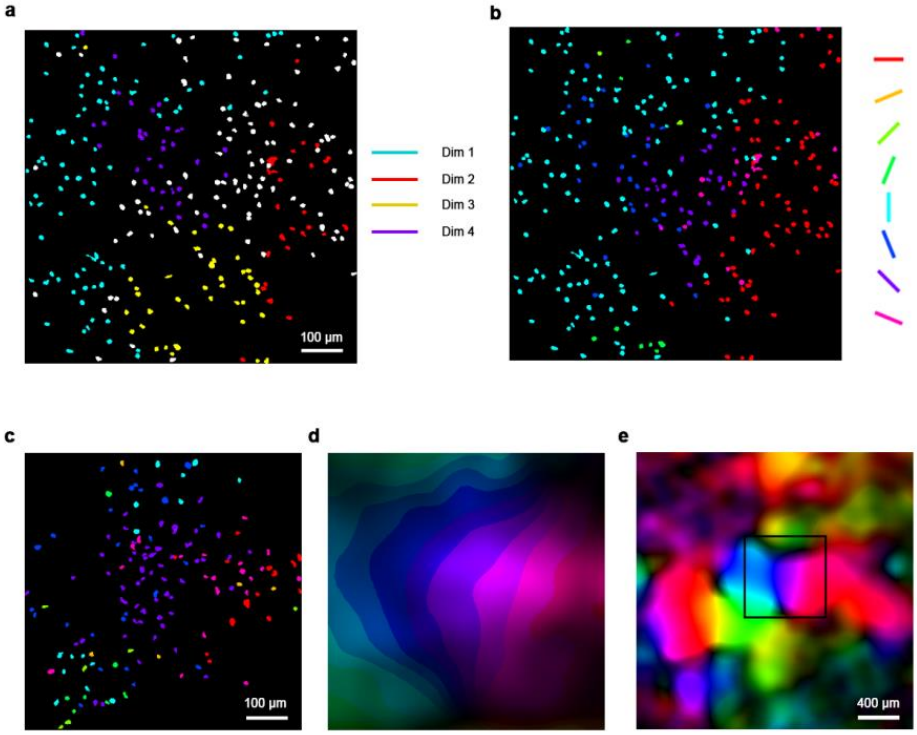

Figure 4: Neuronal orientation selectivity derived via Fourier transformation is consistent with MDS results. (**a**) Cell map of preferred feature dimensions. Each neuron was labelled according to the feature dimension on which its maximum RTA was found. (**b**) Cell map for orientation, derived by applying a Fourier transform to the natural images to extract the dominant orientation. (**c**) Cell map of orientation derived by orientated bar stimuli. (**d**) Filtered map of orientation derived by orientated bar stimuli. Only orientation selective neurons (OSI > 0.3, see Online Methods) were processed. (**e**) Filtered map of orientation derived by orientated bar stimuli at larger scale (4x objective). The recording site for 16x objective is marked by the black box.

Feature dimensions and grating selectivity

Apart from orientation, previous research found some V4 neurons that respond preferentially to concentric gratings or radial gratings, and we wondered how this may be related to the opposing representation of curves and corners we found by MDS. We therefore recorded neuronal responses to Cartesian and non-Cartesian gratings, in the same imaging site. We found some neurons exhibiting high selectivity to concentric gratings of various spatial frequencies and phases while not responding to Cartesian or radial gratings (**Fig. 5a**). For monkey A and monkey B, neurons preferring concentric gratings over radial gratings were spatially clustered in the curvature domains (**Fig. 5b-c**), consistent with dimension 3 in MDS (**Fig. 3f** and **Fig. 5d**). The curvature domains were co-organized with orientation domains, but in a distinctly different structure (**Fig. 5e**). We also examined the relationship between

curvature selectivity and spatial frequency (SF). We found that even though the two monkeys had different optimal spatial frequencies at their recording sites (**Supplementary Fig. 5a-b**), their maps of concentric gratings preference both remained consistent across various spatial frequencies (**Supplementary Fig. 5c**). Therefore, the curvature dimension in V4 is responsible for concentric grating and radial grating selectivity that is not necessarily related to low-order orientation and spatial frequency.

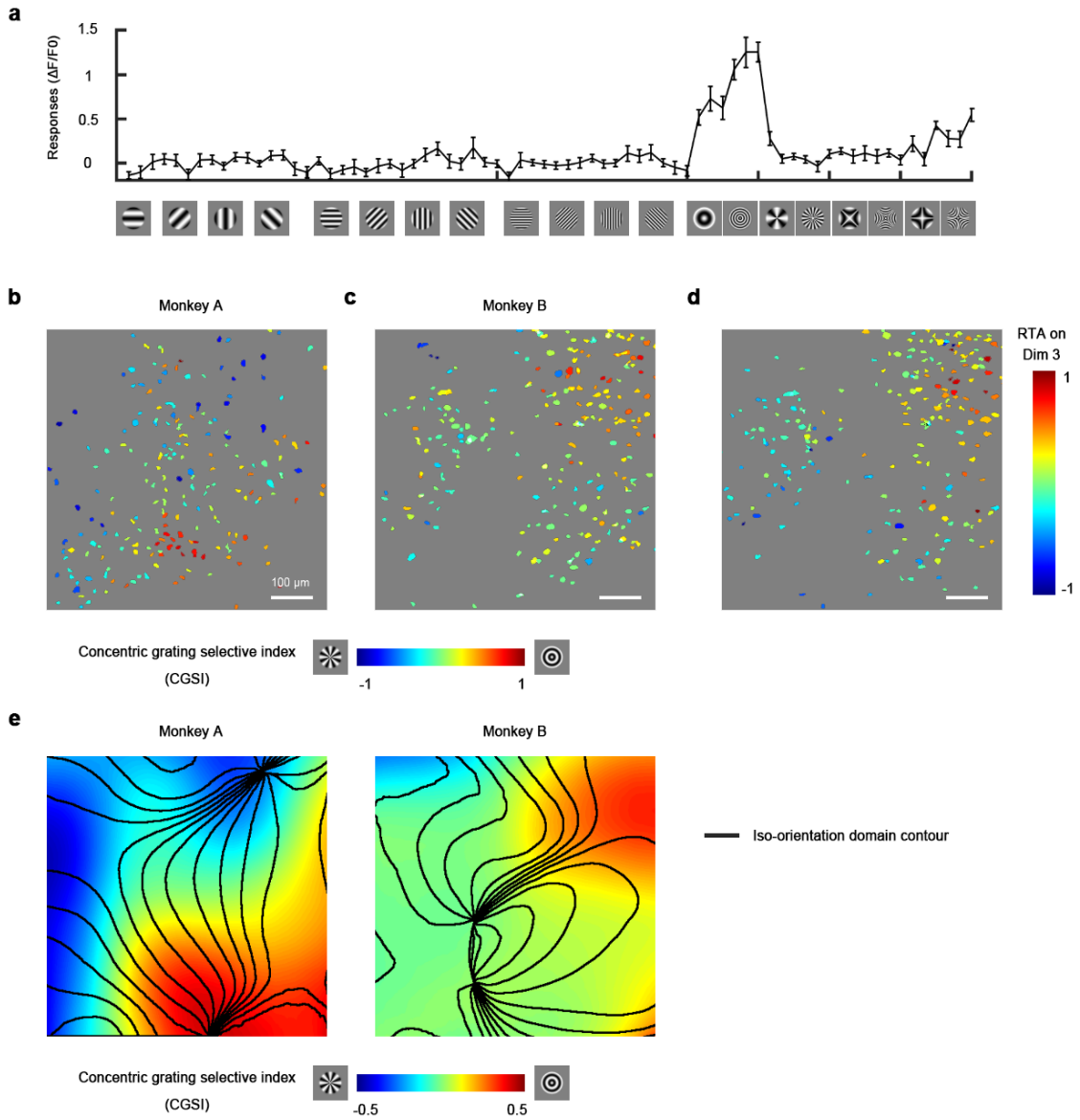

Figure 5: Feature dimensions and grating stimuli selectivity. (**a**) An example neuron that is selective for concentric gratings while not responding to Cartesian or radial gratings. (**b**) Cell map of CGSI (concentric grating selective index, which characterizes a neuron's preference to concentric gratings over radial gratings). Concentric gratings preferring neurons localized in the lower middle of the imaging area, consistent with dimension 3 (curvature dimension). (**c**) Cell map of CGSI for monkey B. (**d**) Cell map of RTA on dimension 3 for monkey B. (**e**) Filtered map of combined organization of orientation and CGSI (curvature and corner).

Curvature and contour feature tuning

To further probe the independence of the curvature dimension, and to investigate neuronal coding for curves and corners in detail, we then conducted experiments to measure neuronal responses to contour feature stimuli at the same recording sites. The stimulus set consisted of 26 forms including bars, corners, polygonal lines and curves, each with 8 orientations.

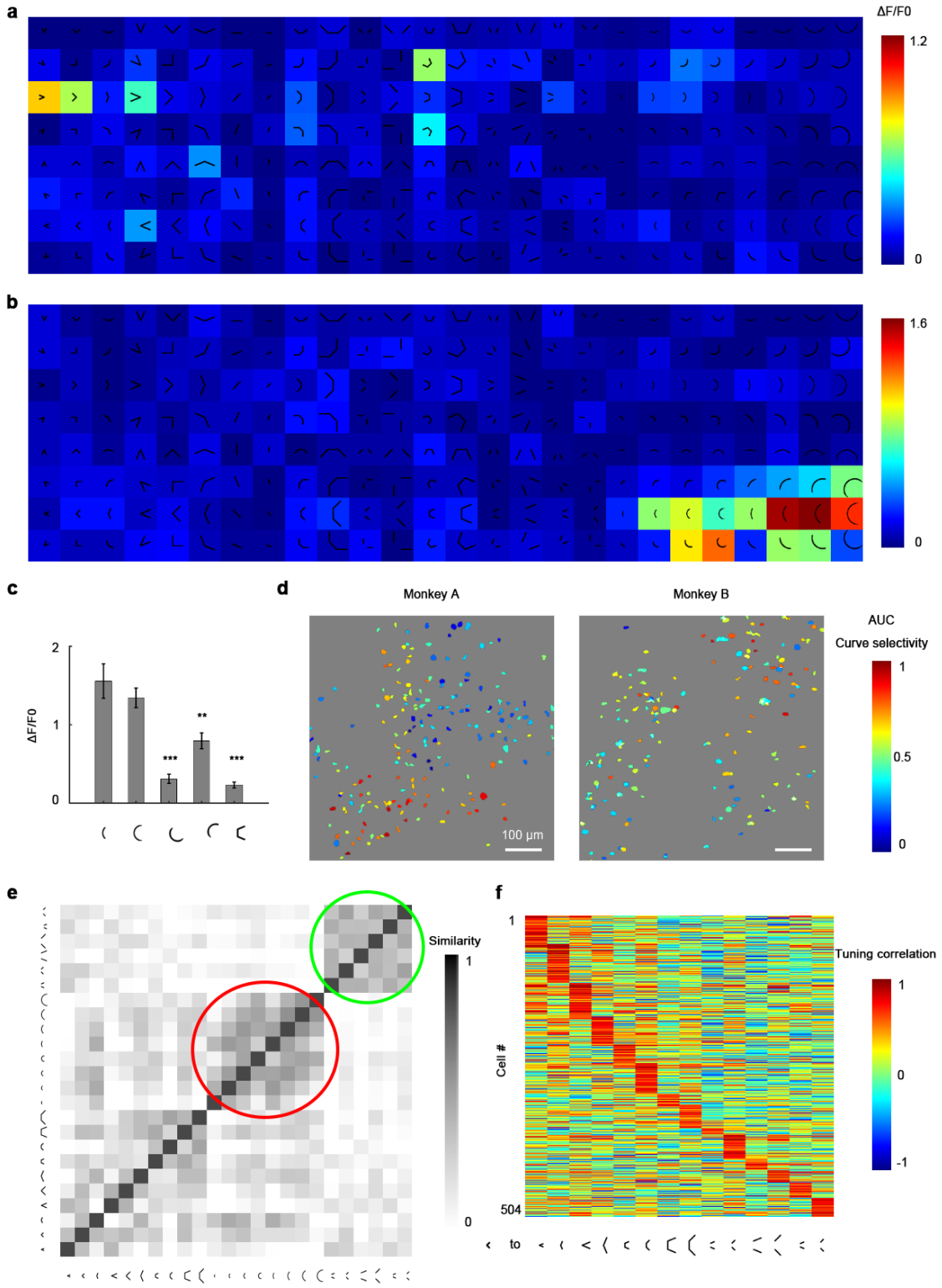

Figure 6 Validation of the existence and independence of curvature/corner dimension by synthetic contour feature stimuli. (**a**) An example neuron selective for corners and polygonal lines. (**b**) An example neuron selective for smooth curvature while not responding to "Π shape" polygonal lines that highly resemble curves. The neuron was also selective to the orientation of the curves. (**c**) Neuronal response remained constant when the 90° curve was symmetrical elongated, but decreased significantly when it was asymmetrical elongated even though the original 90° remained intact, indicating the neuron was tuned to the global orientation of the entire curve instead of local features. (**d)** ROC analysis was performed to quantify neurons' curvature selectivity. Neurons with high AUC (curvature selectivity) were spatially clustered in the curvature patches. (**e**) The simplified similarity matrix derived by population responses showing the similarities of population coding between various stimulus forms. Curve stimuli formed a cluster (red circle). Disjoint bars at various separation angles were also more similar to each other (green circle). (**f**) Neurons' tuning correlation between 90° corner and various stimulus forms. For many of the neurons, tuning to 90° corner was more correlated to 45° and 135° corners than to two disjoint bars perpendicular, indicating the neuron was tuned to the global orientation of the entire corner instead of combination of multiple orientations.

As has been reported previously, many cells exhibited clear preferences to specific features like corners or curves. An example neuron that was selective for corners, but exhibited no curve responses is shown in **Fig. 6a**. Other neurons were highly selective for smooth curvature while not responding to 'Π shape' polygonal lines even though they resembled curves (**Fig. 6b**). Some curvature neurons were also found to be selective to the orientation of the curves (**Fig. 6b** and **Supplementary Fig. 6a**), so we wondered whether it was the local orientation or the orientation of the entire feature that the neurons were selected to. We found neuronal responses remained consistent when the 90° curve at its optimal orientation was elongated symmetrically to 180°, but decreased significantly when it was elongated asymmetrically in which case it still contained the original 90° components but the orientation of the entire curve was changed (**Fig. 6c**). This could also be interpreted as the orientation tuning bandwidth of 180° curve, and we found that the bandwidth (population median) remained narrow when the angle rose to 180°, indicating neurons were selective to the orientation of the curve as an integral feature.

We then examined the clustering of curvature neurons. To quantify neurons' curvature selectivity, we performed ROC analysis to evaluate neurons' performance as a curvature detecting classifier. We found that for both monkeys, neurons with high AUC — smooth curvature selectivity — were spatially clustered, in the same domains revealed by MDS and grating stimuli (**Fig. 6d**), thus confirming the curvature domains. We also evaluated the neural population representation of stimuli by constructing a similarity matrix characterizing the correlation of population responses between various contour feature stimuli (**Fig. 6e** and **Supplementary Fig. 7a**). We found that smooth curves were also clustered by population responses (**Fig. 6e**).

Interestingly, we also found another cluster in the similarity matrix consisting of disjoint bars with different separation angles, sizes and locations (**Fig. 6e**), which were comparatively less similar to corners with the same separation angles. When we examined neurons' tuning correlation between corners and disjoint bars, we found neuronal tuning to 90° corner was more correlated to 45° and 135° corners rather than two disjoint perpendicular bars (**Fig. 6f**). Some corner selective neurons showed the same orientation preference to corners with various separation angles and did not even respond to disjoint bars (**Supplementary Fig. 7b**). This suggested that corners were encoded as an integral feature differently from the two disjoint orientated bars.

Our results with contour feature stimuli, while confirming the existence and spatial clustering of curvature and corner selective neurons, further reveals that the neuronal tuning and population representation of curves and corners cannot be fully explained by local orientation or combination of multiple orientations. We may therefore regard this as a more complex pattern encoding feature dimension in shape representation that is largely independent of low-order orientation.

DISCUSSION

As shape information is processed along the ventral pathway, low-order local orientation information represented in V1[3,4] is transformed and integrated into representation of high-order complex natural objects, scenes, and even faces in the inferotemporal cortex[9,17]. As an intermediate stage of this pathway, V4 could play an important role in this transformation and integration. Early studies demonstrated V4 neurons are tuned to orientations and colors[17,43,44], but the existence of these tunings does not necessarily mean the physiological function of these neurons is merely detecting orientations or colours, and not encoding other more complex information. Other pioneering studies suggested that V4 neurons are selective for more complex features such as corners or curves[21,22]. Whereas predefined parameterized stimulus sets can help to reveal specific details of neural tuning, the fact that it relies on a priori knowledge raises the concern about whether it can objectively reflect the neurons' true physiological function when viewing scenes and objects in natural environments.

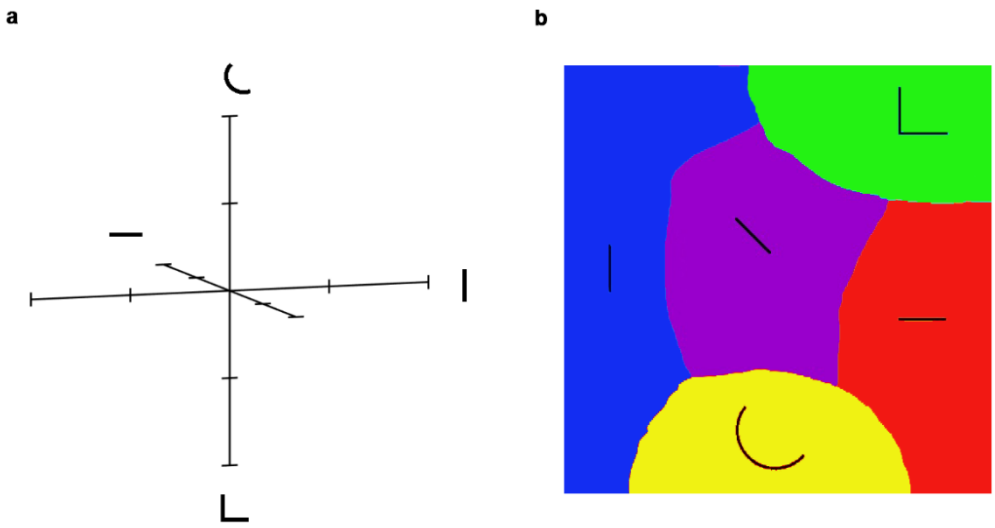

Figure 7 A hypothetical principle for shape representation in V4. (**a**) The major feature dimensions for shape representation were orientations and curvature/corner. Curvature and corner are represented in opposition to one another, orthogonal to the orientation representations. (**b**) These feature dimensions are encoded by spatially clustered subpopulations of neurons organized in iso-orientation domains, curvature domains and corner biased domains.

Here, by using an unbiased stimulus set composed of thousands of natural images and unbiased data-driven categorization solely based on population responses, we demonstrate a feature space for shape representation dominated by orientation, curvature and corner dimensions (**Fig. 7a**). The opponent representation of curves and corners in dimension 3 is consistent with previous studies[18,21]. Tuning to integral feature orientation instead of local compartments predicted by classical endstopping[45,46] is also consistent with the invariant complex feature encoding previously suggested[47,48]. The independence of the curvature dimension from the orientation dimensions implies that simple and complex shape features are both encoded in V4 while viewing natural scenes, and of comparable importance.

These feature dimensions are encoded by different neuron clusters organized in separate functional domains for orientation, curves and corners (**Fig. 7a**). Functional organization of low-order orientation and spatial frequency representation in macaque V4 had been visualized by intrinsic signal optical imaging[15,16], consistent with our results. For more complex features, few studies have been carried out in V4 to characterize the functional organization, let alone at single cellular resolution. We report here cortical micro-domains consisting almost entirely of neurons selective for curves, which can be regarded as curvature domains. This finding at the single cell level is consistent with a previous fMRI study reporting curvature-biased patches in macaque V4[20]. But, to the best of our knowledge, this is the first

time the combined functional organization of simple and complex shape features in V4 has been demonstrated at single cellular resolution.

The co-existence and co-organization of simple and complex shape feature representation in V4 could be of certain functional implication. Neurons in V1 are tuned to low-order orientation and spatial frequency and organized in iso-orientation domains and orientation pinwheels[3,4,49]. Neurons in IT are selective for complex features and objects and organized in feature columns and face patches[10,12,13]. The simple-to-complex transformation and integration take place in the intermediate stages. Researchers reported that some V2 neurons are selective to combination of multiple local orientations, from which corner selectivity might emerge[5,6]. Our results in V4 showed that simple and complex shape features are encoded by different subpopulation of neurons in different functional domains. The curvature or corner selective neurons exhibited a higher degree of complexity compared to V2 as they are tuned to the integral feature instead of the combination of local orientation compartments. This suggests something more complex than a purely feed forward mechanism. It is possible that these complex feature selective neurons also received recurrent modulation from other neurons in the same curvature/corner domain, as well as from nearby orientation domains, which might support the previous finding that the response profile of V4 neurons were temporally heterogeneous[30,50].

Furthermore, our results may also help to explore the later stage of visual hierarchy. Our results suggest that higher order pattern domains may emerge gradually along ventral pathway. The curvature and corner domains in V4 could possibly form the basis of the more complex feature columns and even face patches in IT, and are therefore presumably useful to reveal the neural process of complex objects recognition and classification.

ACKNOWLEDGEMENTS

We thank the Peking University Laboratory Animal Center for animal care. We acknowledge the Genetically Encoded Calcium Indicator (GECI) project at Janelia Farm Research Campus Howard Hughes Medical Institute. We thank Weibin Song, Ziqing Li and Tianye Wang for helping with the experiments, Yang Li and Hongfei Jiang for assistance with surgeries and viral injections, and Ian Max Andolina and Niall Mcloughlin for their comments and suggestion on the manuscript. This work was supported by National Natural Science Foundation of China (grant no. 31730109), National Basic Research Program of China (grant no. 2017YFA0105201), National Natural Science Foundation of China Outstanding Young Researcher Award (grant no. 30525016) and a Project 985 grant of Peking University, Beijing Municipal Commission of Science and Technology (grant no. Z151100000915070).

AUTHOR CONTRIBUTION

S.T. and R.J. conceived the project and designed the experiments. R.J. and M.L. performed the experiments. R.J. analyzed the data and wrote the manuscript. S.T. supervised the project and revised the manuscript.

REFERENCES

1 Mishkin, M., G. Ungerleider, L. & A. Macko, K. Object vision and spatial vision: Two cortical pathways. *Trends in Neurosciences* **6**, 414-417, doi:10.1016/0166-2236(83)90190-X (1983).

2 Connor, C. E., Brincat, S. L. & Pasupathy, A. Transformation of shape information in the ventral pathway. *Curr Opin Neurobiol* **17**, 140-147, doi:10.1016/j.conb.2007.03.002 (2007).

3 Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture of monkey striate cortex. *J Physiol* **195**, 215-243, doi:10.1113/jphysiol.1968.sp008455 (1968).

4 Hubel, D. H. & Livingstone, M. Segregation of form, color, and stereopsis in primate area 18. *Journal of Neuroscience* **7**, 3378-3415 (1987).

5 Ito, M. & Komatsu, H. Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *J Neurosci* **24**, 3313-3324, doi:10.1523/JNEUROSCI.4364-03.2004 (2004).

6 Anzai, A., Peng, X. & Van Essen, D. C. Neurons in monkey visual area V2 encode combinations of orientations. *Nat Neurosci* **10**, 1313-1321, doi:10.1038/nn1975 (2007).

7 Ts'o, D. Y., Frostig, R. D., Lieke, E. E. & Grinvald, A., . Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* **249**, 417-420 (1990).

8 Desimone, R., Albright, T. D., Gross, C. G. & Bruce, C. Stimulus-selective properties of inferior temporal neurons in the macaque. *The Journal of neuroscience : the official journal of the Society for Neuroscience* **4**, 2051-2062, doi:10.1523/jneurosci.04-08-02051.1984 (1984).

9 Fujita, I., Tanaka, K., Ito, M. & Cheng, K. Columns for visual features of objects in monkey inferotemporal cortex. *Nature* **360**, 343-346 (1992).

10 Tsao, D. Y., Freiwald, W. A., Tootell, R. B. & Livingstone, M. S. A cortical region consisting entirely of face-selective cells. *Science* **311**, 670-674, doi:10.1126/science.1119983 (2006).

11 Freiwald, W. A. & Tsao, D. Y. Functional compartmentalization and viewpoint generalization within the macaque face-processing system. *Science* **330**, 845-851, doi:10.1126/science.1194908 (2010).

12 Tsunoda, K., Yamane, Y., Nishizaki, M. & Tanifuji, M. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nature Neuroscience* **4**, 832-838 (2001).

13 Tanaka, K. Columns for complex visual object features in the inferotemporal cortex: clustering of cells with similar but slightly different stimulus selectivities. *Cereb Cortex* **13**, 90-99, doi:10.1093/cercor/13.1.90 (2003).

14 Roe, A. W. *et al.* Toward a unified theory of visual area V4. *Neuron* **74**, 12-29, doi:10.1016/j.neuron.2012.03.011 (2012).

15 Tanigawa, H., Lu, H. D. & Roe, A. W. Functional organization for color and orientation in macaque V4. *Nat Neurosci* **13**, 1542-1548, doi:10.1038/nn.2676 (2010).

16 Lu, Y. *et al.* Revealing Detail along the Visual Hierarchy: Neural Clustering Preserves Acuity from V1 to V4. *Neuron* **98**, 417-428 e413, doi:10.1016/j.neuron.2018.03.009 (2018).

17 Desimone, R. & Schein, S. J. Visual properties of neurons in area V4 of the macaque: sensitivity to stimulus form. *Journal of Neurophysiology* **57**, 835-868 (1987).

18 Gallant, J., Braun, J. & Van Essen, D. Selectivity for polar, hyperbolic, and Cartesian gratings in macaque visual cortex. *Science* **259**, 100-103 (1993).

19 Kobatake, E. & Tanaka, K. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology* **71**, 856-867 (1994).

20 Yue, X., Pourladian, I. S., Tootell, R. B. & Ungerleider, L. G. Curvature-processing network in macaque visual cortex. *Proc Natl Acad Sci U S A* **111**, E3467-3475, doi:10.1073/pnas.1412616111 (2014).

21 Pasupathy, A. & Connor, C. E. Responses to Contour Features in Macaque Area V4. *Journal of Neurophysiology* **82**, 2490 (1999).

22 Pasupathy, A. & Connor, C. E. Shape Representation in Area V4: Position-Specific Tuning for Boundary Conformation. *Journal of Neurophysiology* **86**, 2505-2519 (2001).

23 Pasupathy, A. & Connor, C. E. Population coding of shape in area V4. *Nat Neurosci* **5**, 1332-1338, doi:10.1038/nn972 (2002).

24 Cadieu, C. *et al.* A model of V4 shape selectivity and invariance. *J Neurophysiol* **98**, 1733-1750, doi:10.1152/jn.01265.2006 (2007).

25 Bushnell, B. N., Harding, P. J., Kosai, Y. & Pasupathy, A. Partial occlusion modulates contour-based shape encoding in primate area V4. *J Neurosci* **31**, 4012-4024, doi:10.1523/JNEUROSCI.4766-10.2011 (2011).

26 Carlson, E. T., Rasquinha, R. J., Zhang, K. & Connor, C. E. A sparse object coding scheme in area V4. *Curr Biol* **21**, 288-293, doi:10.1016/j.cub.2011.01.013 (2011).

27 Okazawa, G., Tajima, S. & Komatsu, H. Image statistics underlying natural texture selectivity of neurons in macaque V4. *Proc Natl Acad Sci U S A* **112**, E351-360, doi:10.1073/pnas.1415146112 (2015).

28 Ziemba, C. M., Freeman, J., Movshon, J. A. & Simoncelli, E. P. Selectivity and tolerance for visual texture in macaque V2. *Proc Natl Acad Sci U S A* **113**, E3140-3149, doi:10.1073/pnas.1510847113 (2016).

29 Nandy, A. S., Sharpee, T. O., Reynolds, J. H. & Mitchell, J. F. The fine structure of shape tuning in area V4. *Neuron* **78**, 1102-1115, doi:10.1016/j.neuron.2013.04.016 (2013).

30 Nandy, A. S., Mitchell, J. F., Jadi, M. P. & Reynolds, J. H. Neurons in Macaque Area V4 Are Tuned for Complex Spatio-Temporal Patterns. *Neuron* **91**, 920-930, doi:10.1016/j.neuron.2016.07.026 (2016).

31 Vinje, W. E. & Gallant, J. L. Sparse Coding and Decorrelation in Primary Visual Cortex During Natural Vision. *Science* **287**, 1273-1276 (2000).

32 Ringach, D. L., J., H. M. & Robert, S. Receptive field structure of neurons in monkey primary visual cortex revealed by stimulation with natural image sequences. *J Vis* **2**, 12-24 (2002).

33 Willmore, B. D., Prenger, R. J. & Gallant, J. L. Neural representation of natural images in visual area V2. *J Neurosci* **30**, 2102-2114, doi:10.1523/JNEUROSCI.4099-09.2010 (2010).

34 Tang, S. *et al.* Large-scale two-photon imaging revealed super-sparse population codes in the V1 superficial layer of awake monkeys. *Elife* **7**, doi:10.7554/eLife.33370 (2018).

35 Sharpee, T. O., Kouh, M. & Reynolds, J. H. Trade-off between curvature tuning and position invariance in visual area V4. *Proc Natl Acad Sci U S A* **110**, 11618-11623, doi:10.1073/pnas.1217479110 (2013).

36 Nauhaus, I., Nielsen, K. J., Disney, A. A. & Callaway, E. M. Orthogonal micro-organization of orientation and spatial frequency in primate primary visual cortex. *Nat Neurosci* **15**, 1683-1690, doi:10.1038/nn.3255 (2012).

37 Li, M., Liu, F., Jiang, H., Lee, T. S. & Tang, S. Long-Term Two-Photon Imaging in Awake Macaque Monkey. *Neuron* **93**, 1049-1057 e1043, doi:10.1016/j.neuron.2017.01.027 (2017).

38 Garg, A. K., Li, P., Rashid, M. S. & Callaway, E. M. Color and orientation are jointly coded and spatially organized in primate primary visual cortex. *Science* **364**, 1275-1279, doi:10.1126/science.aaw5868 (2019).

39 Chang, L. & Tsao, D. Y. The Code for Facial Identity in the Primate Brain. *Cell* **169**, 1013-1028 e1014, doi:10.1016/j.cell.2017.05.011 (2017).

40 Freiwald, W. A., Tsao, D. Y. & Livingstone, M. S. A face feature space in the macaque temporal lobe. *Nat Neurosci* **12**, 1187-1196, doi:10.1038/nn.2363 (2009).

41 Hegde, J. & Van Essen, D. C. A comparative study of shape representation in macaque visual areas V2 and V4. *Cereb Cortex* **17**, 1100-1116, doi:10.1093/cercor/bhl020 (2007).

42 Schwartz, O., Pillow, J. W., Rust, N. C. & Simoncelli, E. P. Spike-triggered neural characterization. *Journal of Vision* **6**, 484 (2006).

43 Zeki, S. Colour coding in rhesus monkey prestriate cortex. Brain Res. 27, 422-427. *Brain research* **53**, 422-427, doi:10.1016/0006-8993(73)90227-8 (1973).

44 Conway, B. R., Moeller, S. & Tsao, D. Y. Specialized color modules in macaque extrastriate cortex. *Neuron* **56**, 560-573, doi:10.1016/j.neuron.2007.10.008 (2007).

45 Hubel, D. H. & Wiesel, T. N. Receptive fields and functional architecture in two nonstriate visual areas (18 and 19) of the cat. *Journal of Neurophysiology* **28**, 229-289 (1965).

46 Dobbins, A., Zucker, S. W. & Cynader, M. S. Endstopped neurons in the visual cortex as a substrate for calculating curvature. *Nature* **329**, 438-441 (1987).

47 El-Shamayleh, Y. & Pasupathy, A. Contour Curvature As an Invariant Code for Objects in Visual Area V4. *Journal of Neuroscience* **36**, 5532-5543 (2016).

48 Oleskiw, T. D., Pasupathy, A. & Bair, W. Spectral receptive fields do not explain tuning for boundary curvature in V4. *J. Neurophysiol.* **112**, 2114-2122, doi:10.1152/jn.00250.2014 (2014).

49 Ts'O, D. Y., Frostig, R. D., Lieke, E. E. & Grinvald, A., . Functional organization of primate visual cortex revealed by high resolution optical imaging. *Science* **249**, 417-420 (1990).

50 Yau, J. M., Pasupathy, A., Brincat, S. L. & Connor, C. E. Curvature processing dynamics in macaque area V4. *Cereb Cortex* **23**, 198-209, doi:10.1093/cercor/bhs004 (2013).

METHODS

All procedures involving animals were in accordance with the Guide of Institutional Animal Care and Use Committee (IACUC) of Peking University Animals, and approved by the Peking University Animal Care and Use Committee (LSC-TangSM-5).

Animal preparation. The subjects used in this study were two adult male rhesus monkeys (*Macaca mulatta*, 4 and 5 years of age, respectively). Two sequential surgeries were performed on each animal under general anesthesia. In the first surgery, we performed a craniotomy over V4 and opened the dura. We injected 200 nl of AAV1.Syn.GCaMP6s.WPRE.SV40 (CS0564, titer 2.2e13 [GC/ml], Penn Vector Core) or AAV1.hSyn.GCaMP5G.WPRE.SV40 (V4102MI-R, titer 2.37e13 [GC/ml], Penn Vector Core) at a depth of about 350 µm. Injection and surgical protocols followed our previous study. After injections, we sutured the dura, replaced the skull cap with titanium screws, and closed the scalp. The animal was then returned for recovery and received Ceftriaxone sodium antibiotic (Youcare Pharmaceutical Group Co. Ltd., China) for one week. 45 days later, we performed the second surgery to implant the imaging window and head posts. The dura was removed and a glass coverslip was put directly above the cortex without any artificial dura. The detailed design of the chamber and head posts can be found in our previous study (ref.39).

Visual stimuli. The natural images were downloaded from ImageNet and cropped into 3° round patch. The gratings and contour feature stimuli were generated by Matlab (The MathWorks, Natick, MA). The stimuli were displayed on a LCD monitor 45 cm from the animal's eyes (30 pixels per degree). Each

stimulus was presented for 1 s interleaved by 1.5 s blank, in random order. The RF location was estimated by drifting gratings. The eccentricities were around 1° for both monkeys.

Two-photon imaging. Monkeys were trained to hold fix on a white spot on a LCD monitor for 2.5 s with their heads fixed to receive a juice reward. 1000 nm mode-lock laser was used for excitation of GCaMPs, and resonant scanning (512x512 pixels, 32 frames per second) was used to record the fluorescence images (8 fps, averaging every 4 frames). A 16x objective (850x850 µm) was used for neuronal population recording at single-cellular resolution, and 4x objective for sub-cortical level recording.

Image data was processed by Matlab. The two-photon images were all aligned to a template images to eliminate shift artifacts. The differential fluorescence image was calculated as ΔF = F – F0, where the basal fluorescence image F0 was defined as the average image of 0-0.5 s before stimulus onset, and F as the average of 0.5-1.25 s after stimulus onset. To identify responding cell bodies (ROIs), the differential images went through a band-pass filter and were binarized. The connected components (>25 pixels) were identified as active neurons.

Multidimensional scaling. The distance matrix used for MDS of 4136 natural image stimuli was constructed as

$$D_{ij} = 1 - corr2\left(\Delta F_i, \Delta F_j\right),$$

where ΔF_i was the filtered differential image differential image, and D_{ij} was the distance between stimulus i and j.

Multidimensional scaling was performed by Matlab command *cmdscale*, or *mdscale* with criterion *strain*. The stimuli were represented in a new 20-dimensional space as

$$\overrightarrow{S_i} = \left(Y_{i1}, Y_{i2}, Y_{i3} \ldots Y_{i20}\right),$$

where Y was the 4136 x 20 coordinate matrix, and Y_{ik} was the coordinate of stimulus i on dimension k.

The pairwise distance in the new stimulus space was calculated as the Euclidean distance as

$$N_{ij} = \left\| \overrightarrow{S_i} - \overrightarrow{S_j} \right\| = \sqrt{\sum_k \left(Y_{ik} - Y_{jk}\right)^2},$$

and the normalized stress was defined as

$$Stress = \frac{\sum\sum\left(D_{ij} - N_{ij}\right)^2}{\sum\sum {D_{ij}}^2},$$

which was the total loss of distance after dimensional reduction.

The activation map (M) for dimension k was defined as the weighted average of differential fluorescence images as

$$M_k = \sum_{i=1}^{4136} \Delta F_i \cdot Y_{ik},$$

Where ΔF_i was the differential image for stimulus i.

Response triggered average. The population average response triggered average (PARTA, 1 x 20) was defined as the average of coordinates (normalized by 3x S.D. for each dimension) weighted by population average responses:

$$\overrightarrow{PARTA} = \overrightarrow{Ra} * Y \,/\, \sum Ra\ ,$$

where Ra was the population average response to 4136 stimuli (1 x 4136), and Y was the coordinate matrix (4136 x 20). PARTA > 0 on a dimension if more neurons responded strongly to stimuli that had positive coordinate. The polarity of each axis of feature dimension was re-defined according to the PARTA. If PARTA > 0 on a dimension, we reversed the polarity so that the stimuli were sparsely represented in the forward direction.

The response triggered average of a single neuron was defined as

$$\overrightarrow{RTA} = \left(\overrightarrow{Rn} * Y \,/\, \sum Rn - \overrightarrow{PARTA}\right) .* \ sign\left(-\overrightarrow{PARTA}\right),$$

where Rn was a single neuron's responses to 4136 stimuli (1 x 4136). The *signum* function was used to reverse the polarity of axes whose PARTA > 0. High RTA value on a dimension means the neuron is selective to the stimuli with large coordinates.

Fourier transform. We performed Fourier transform to extract the orientation information form natural image stimuli after edge detection, so that each stimuli was decomposed into an orientation

spectrum (OS, 4136 x 180). A neuron's orientation tuning (OT, 1 x 180) was defined as its average responses to 180 orientations weighted by the orientation spectrums:

$$\overrightarrow{OT} = \overrightarrow{Rn} * OS \,./\, OSsum\,,$$

where Rn was a single neuron's responses to 4136 stimuli (1 x 4136), and OSsum (1 x 180) was the summation of all spectrums.

Orientation selective index. The orientation selective index for bars was defined as

$$OSI = \frac{resp(optimal\ orientation) - resp(optimal\ orientation{+}90°)}{resp(optimal\ orientation) + resp(optimal\ orientation{+}90°)}.$$

Neurons whose OSI > 0.3 were regarded as orientation selective neurons.

Orientation selective index for curves was defined as

$$OSI = \frac{resp(optimal\ orientation) - resp(optimal\ orientation{+}180°)}{resp(optimal\ orientation) + resp(optimal\ orientation{+}180°)},$$

where resp (optimal orientation) was the neuron's maximum response to curves with 0.67° radius and 180° radian.

Orientation map. We first applied a Gaussian filter to the differential fluorescence images. We then calculated the vector summation of responses (ΔF/F0) to 8 orientations for each pixel. We assigned a hue to each pixel according to the orientation of the sum vector, and brightness according to the vector length, in which way a pseudo color orientation map was made.

Concentric grating selective index. Concentric selective index was defined as

$$CGSI = \frac{meanresp(concentric) - meanresp(radial)}{meanresp(concentric) + meanresp(radial)}.$$

ROC analysis. ROC analysis was performed to quantify Neurons' selectivity to curve stimuli. Neurons were presumed to be a binary classifier to distinguish curves. For each neuron, 26 instances in total were

considered, each being the neuron's maximum response to a stimulus form among 8 orientations. The 8 curve stimuli were labeled as '1', and others as '0'.

Tuning bandwidth. Neurons' bar orientation tuning was fitted to Gaussian function as

$$\mathrm{r}(\theta) = A \cdot exp\left(-\frac{(\theta-\mu)^2}{2\sigma^2}\right) + B,$$

where μ was the optimal orientation and σ was the tuning half-bandwidth.

Curve orientation tuning was fitted in a similar way as

$$\mathrm{r}(\theta) = A_1 \cdot exp\left(-\frac{(\theta-\mu)^2}{2{\sigma_1}^2}\right) + A_2 \cdot exp\left(-\frac{(\theta-\mu+\pi)^2}{2{\sigma_2}^2}\right) +$$

$$A_2 \cdot exp\left(-\frac{(\theta-\mu-\pi)^2}{2{\sigma_2}^2}\right) + B,$$

where μ was the optimal orientation and σ_1 was the tuning half-bandwidth.

Neurons' whose fitting $R^2 > 0.5$ were used to calculate the median.

Similarity matrix. The similarity matrix was used to compare the population coding for contour features and bars. The population response vector used to construct the matrix was defined as

$$\overrightarrow{PR_i} = \left(R_{1i},\ R_{2i},\ R_{3i}\ \ldots\right),$$

where R_{ki} was the response of neuron k to stimulus i. Each element in the 192x192 similarity matrix (24 forms x 8 orientations, 16 single bars out of 208 stimuli were excluded) was defined as the Pearson correlation of two population response vectors. The paradiagonal stripes indicates similar tuning between two forms.

The 192x192 similarity matrix can be separated into 24x24 submatrices, so that each 8x8 submatrix characterizes a pair of stimulus forms. We simplified the similarity matrix to a new 24x24 matrix, in which each element was derived by the trace of the 8x8 submatrix, as

$$SS_{ij} = trace(S(8i-7{:}8i, 8j-7{:}8j))/8 - mean(S(:)),$$

where S was the 192x192 similarity matrix and SS was the 24x24 simplified matrix.