Spatial structure of neuronal receptive field in awake monkey secondary visual cortex (V2)

Lu Liu[a,b,1], Liang She[a,b,1], Ming Chen[a,b], Tianyi Liu[a], Haidong D. Lu[a,2], Yang Dan[c], and Mu-ming Poo[a,3]

[a]Institute of Neuroscience, State Key Laboratory of Neuroscience, Key Laboratory of Primate Neurobiology, CAS Center for Excellence in Brain Science and Intelligence Technology, Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200031, China; [b]University of Chinese Academy of Sciences, Shanghai 200031, China; and [c]Division of Neurobiology, Department of Molecular and Cell Biology, Howard Hughes Medical Institute, Helen Wills Neuroscience Institute, University of California, Berkeley, CA 94720

Contributed by Mu-ming Poo, January 6, 2016 (sent for review October 12, 2015; reviewed by Judith Hirsch and Doris Y. Tsao)

Visual processing depends critically on the receptive field (RF) properties of visual neurons. However, comprehensive characterization of RFs beyond the primary visual cortex (V1) remains a challenge. Here we report fine RF structures in secondary visual cortex (V2) of awake macaque monkeys, identified through a projection pursuit regression analysis of neuronal responses to natural images. We found that V2 RFs could be broadly classified as V1-like (typical Gabor-shaped subunits), ultralong (subunits with high aspect ratios), or complex-shaped (subunits with multiple oriented components). Furthermore, single-unit recordings from functional domains identified by intrinsic optical imaging showed that neurons with ultralong RFs were primarily localized within pale stripes, whereas neurons with complex-shaped RFs were more concentrated in thin stripes. Thus, by combining single-unit recording with optical imaging and a computational approach, we identified RF subunits underlying spatial feature selectivity of V2 neurons and demonstrated the functional organization of these RF properties.

receptive field | V2 | natural images | stripes

Visual perception depends on processing of the input signals through multiple stages of the visual pathway. At each stage of processing, neuronal representation of the input is governed by the receptive field (RF) properties of individual cells. Characterization of neuronal RFs is a crucial step for understanding the mechanism of visual processing.

Compared with the primary visual cortex (V1), neuronal RFs in the secondary visual cortex (V2) are much less understood. Previous studies have shown that in addition to orientation and direction selectivity (1, 2), similar to that found in V1, neurons in V2 also exhibit selectivity for more complex spatial features such as angle (3–5), illusory contour (6), complex shapes (7), texture (8), and segmentation of the scene (9, 10). Given the large number of potentially relevant visual features, traditional methods using stimulus sets with parametric variation of particular visual features are not efficient for comprehensive RF characterization.

An alternative approach is to fit the stimulus–response relationship of each neuron by a parametric model. The relationship is ideally probed with large ensembles of visual stimuli, and the resulting model can be used to predict the neuronal responses to other arbitrary stimuli (11, 12). Natural stimuli are well suited for this purpose, because the visual system has evolved to process natural scenes, which contain rich spatial features that are more effective than random stimuli in eliciting cortical responses (13). Such an approach imposes no prior assumption about which stimulus features are relevant to the cell and is thus well suited for unbiased RF characterization.

In the present study, we used large ensembles of natural images to probe the neuronal responses in awake macaque monkeys and a linear–nonlinear model (14, 15) to represent the RF of each V2 neuron. The subunits of the RF models were identified by a method adapted from projection pursuit regression (PPR) (16–18), which does not require stimuli with specific statistical properties and is thus well suited for analyzing the neuronal responses to natural stimuli. Compared with other optimization methods, a distinct feature of PPR is to optimize one subunit of the RF model at a time to reduce the dimensionality of the problem. Using this method, we revealed the spatial RF structures of many V2 neurons. Furthermore, we identified cytochrome oxidase stripes of V2 by optical imaging of intrinsic signals. Data from single-unit recording were used in combination with the latter information to determine the spatial organization of cells with different RF properties with respect to the stripes of V2.

Results

Analysis of RF Subunits and Performance of the RF Model. We recorded well-isolated single-unit responses of V2 ($n = 360$) and V1 ($n = 124$) neurons from four hemispheres of three macaque monkeys. The recording was made while the monkey was performing a simple fixation task, and a large ensemble of temporally random grayscale images drawn from a natural scene database (19) were presented in a region covering the RF of the recorded neuron at a rate of 20 images per second (Fig. 1*A* and *Materials and Methods*). We used a linear–nonlinear model for quantitative description of the RF for each neuron (Fig. 1*B*). In the RF model, the firing rate was determined by passing each image (S) through a set of linear filters (F, subunits), and the output of each filter was transformed through a static nonlinear function (f) before summation. The spatial structure of each subunit, its associated static nonlinear function, and the number of subunits were optimized by the PPR algorithm based on the recorded responses to natural images (*Materials and Methods*).

Significance

Using a computational method to analyze neuronal responses evoked by natural scene stimuli, we performed a comprehensive identification of secondary visual cortex (V2) neuronal receptive fields (RFs) and found several novel spatial structures of RFs. This approach imposes no assumption about the selectivity of the neurons, thus allowing a more objective search of RFs. Furthermore, by combining single-unit recording with optical imaging of intrinsic signals, we examined the spatial distribution of V2 neurons exhibiting different RFs, with respect to the V2 stripes defined previously by cytochrome oxidase staining. The identified V2 RFs could be explained by convergence of V1 neurons with well-known primary visual cortex RFs. This study illustrates that computational approach is useful for comprehensive RF identification in higher visual cortices.

Author contributions: L.L., L.S., H.D.L., Y.D., and M.-m.P. designed research; L.L., L.S., M.C., and T.L. performed research; L.L., L.S., and M.C. analyzed data; L.L., L.S., Y.D., and M.-m.P. wrote the paper; and T.L. trained the monkey.

Reviewers: J.H., University of Southern California; and D.Y.T., California Institute of Technology.

The authors declare no conflict of interest.

[1]L.L. and L.S. contributed equally to this work.

[2]Present address: State Key Laboratory of Cognitive Neuroscience and Learning & IDG/McGovern Institute for Brain Research, Beijing Normal University, Beijing 100875, China.

[3]To whom correspondence should be addressed. Email: mpoo@ion.ac.cn.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1525505113/-/DCSupplemental.

NEUROSCIENCE

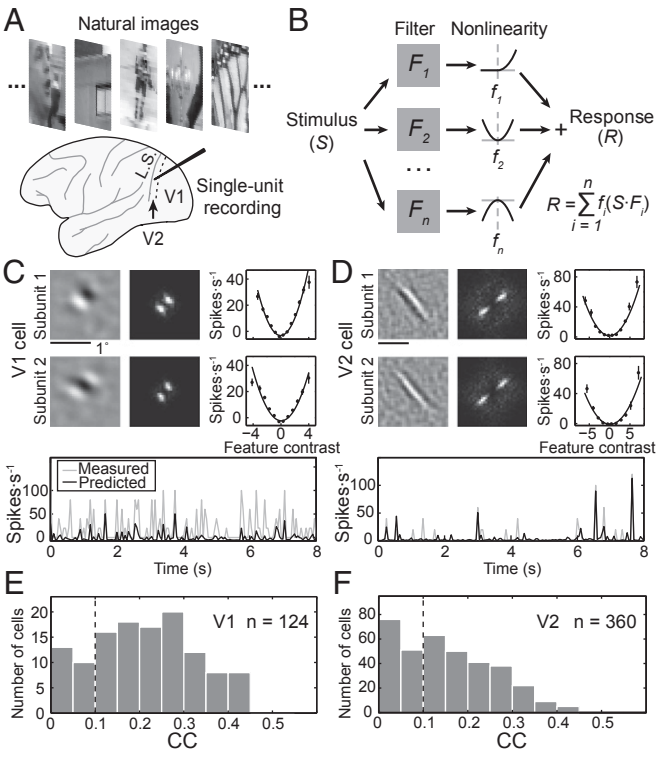

Fig. 1. Analysis of RF subunits. (*A*) Random sequences of natural images were presented and single-unit recordings were made in either V1 or V2. L.S., lunate sulcus; dashed line, V1/V2 border. (*B*) Linear-nonlinear RF model. Each subunit is represented by a linear filter (F_i) depicting the spatial RF structure. *S*, stimulus. The response ($x_i = S{\cdot}F_i$) of each linear filter is passed through a second-order polynomial $f_i(x_i) = a_i + b_i x_i + c_i x_i^2$ before summation. *n*, number of subunits; *i*, the *i*th subunit; *R*, neuronal response. (*C* and *D*) Prediction of responses by RF models of example V1 (*C*) and V2 (*D*) cells. (*Upper*) *Left*, spatial RF; *Center*, spatial spectrum of RF; *Right*, static nonlinear function, with data points (mean ± SEM) depicting measured responses and curves representing polynomial fits of the data. (*Lower*) Gray, measured responses in test data set; black, model prediction. CC between measured (test data) and predicted responses was 0.38 in *C* and 0.37 in *D*. (*E* and *F*) Distribution of CCs for V1 and V2 cells. Black dashed line, threshold for all cells used for further RF analysis.

Examples of the RF models obtained by this analysis for V1 and V2 cells are shown in Fig. 1 *C* and *D* (*Upper*), including the subunits representing the spatial structure of the RF (*Left*), spatial spectrum of the subunits (*Center*), and the corresponding static nonlinear functions (*Right*). We then used the RF model of each cell to predict its responses in a test dataset obtained from the same recording session but not used for computing the RF model (Fig. 1 *C* and *D*, *Lower*). The performance of the model was measured by the correlation coefficient (CC) between the predicted and measured responses, and the distributions of CCs for both V1 and V2 neurons are shown in Fig. 1 *E* and *F*. The PPR method performed significantly better than the spike-triggered average/covariance analysis for natural stimuli in both simulation and experimental data (Fig. S1). For further analyses of the RF structure, only neurons with CC > 0.1 were included (223 V2 and 99 V1 cells).

Spatial Structure of RF Subunits of V2 Cells. Various types of RF subunits found for V1 and V2 cells are shown in Fig. 2*A*, together with the spatial spectrum and associated nonlinear functions. Consistent with previous reports (20–22), we found that for most V1 neurons the RF contained multiple Gabor-like subunits of the same orientation (Fig. 2*A*, cell 2), whereas a small number of them were unoriented with a single center/surround subunit (cell 1), similar to that found for lateral geniculate nucleus (LGN) neurons. However, the RFs of V2 neurons were more diverse. Some of them were "V1-like," with typical Gabor-shaped (cell 4) or unoriented (cell 3) subunits. For other neurons, although the RF subunits can also be modeled as Gabor functions, the ON/OFF subregions were much more elongated than those in typical V1 RFs (cells 5 and 6), and they were referred to as "ultralong" RFs. Furthermore, we found V2 neurons with "complex-shaped" RFs contained multiple preferred orientations (cells 7–9). Diverse complex-shaped RF structures were also reflected by the spatial spectrum of the subunits. As illustrated in Fig. S2*A*, these complex-shaped RF subunits could be decomposed into two Gabor functions with different orientations. These RF subunits may underlie the distinct RF properties previously reported for V2 cells, such as selectivity for combinations of orientations (3–5) and 2D translation invariance to spot stimulus (23, 24).

Quantitative Characterization of RF Structures. To facilitate quantitative comparison between V1 and V2 neurons, we fitted their RF subunits with 2D Gabor functions. First, we fitted each subunit with a single Gabor and computed the aspect ratio (AR), defined as the ratio between the widths (SDs) of the Gaussian envelope along the axes parallel and orthogonal to the sine wave grating (Fig. 2*B*). The AR of each neuron was averaged for all excitatory subunits. We found that the AR distributions were significantly different between V1 and V2 neurons (Fig. 2*B*; $P = 0.006$, Kolmogorov–Smirnov test), with more V2 neurons exhibiting longer RFs (with high ARs).

Second, we fitted each subunit with the sum of two Gabor functions and quantified its complexity by a dual component index (DCI), defined as $R{\bullet}\theta$, where R is the energy ratio between the two Gabor functions (ranged from 0 to 1; Fig. 2*C*), and θ is the angle between the orientations of the two Gabors (ranged 0° to 90°; Fig. 2*C*). Large DCI indicates that the RF consists of at least two Gabors of different orientations but comparable amplitudes. Because center-surround (unoriented) subunits could also be well fitted by two orthogonal Gabor functions with similar amplitude, we identified them by fitting each subunit with a difference of Gaussians (DOG) function; a subunit was classified as center-surround if the goodness-of-fit was >4 (*Materials and Methods*) and excluded from the DCI analysis. The DCI of each neuron was averaged for all excitatory subunits. We found that V2 cells showed significantly higher DCI than V1 cells (Fig. 2*C*; $P = 0.0004$, Kolmogorov–Smirnov test), indicating a higher percentage of complex-shaped RFs.

Evaluation of RF Models. To test whether the ultralong and complex-shaped RFs obtained by PPR are functionally relevant, we tested the performance of the RF model in predicting the neuronal responses (in the test dataset) after manipulating various features of the RF subunits. For cells with ultralong RFs, we first showed that best-fit Gabor functions for the subunits well represented the RFs, with CC close to or slightly better than that of the measured RFs in predicting the neuronal responses (example in Fig. S3*B* and summary in Fig. S3*C*; $P = 0.35$, Wilcoxon signed-rank test). Furthermore, we found that shortening the length of this Gabor function by 60% worsened the prediction of the model, as shown by the increased deviation of the predicted from the measured responses (Fig. S3*A*), especially for natural image stimuli containing long edges aligned with the RF subunits (Fig. S3*A*, *Insets*). Progressive shortening of the Gabor function (from 100% to 40%) resulted in monotonic reduction of the model's performance, as evaluated by CC (example in Fig. S3*B*). The summary of this analysis for all cells with AR >2 showed significant decrease of CC at 60% and 40% length (Fig. S3*C*; 80%, $P = 0.06$; 60%, $P = 0.017$; 40%, $P = 0.0008$; Wilcoxon signed-rank test). Thus, the ultralong RFs significantly contribute to the neuronal responses to natural stimuli.

For cells with complex-shaped RF subunits, we first tested the necessity of each subunit. As shown by an example cell with three subunits in Fig. S4*C*, deleting any one (upper curve) or two (lower curve) of the subunits impaired the model's performance as indicated by the decrease of CC. We also noted that the first subunit contributed much more than the second and third subunits

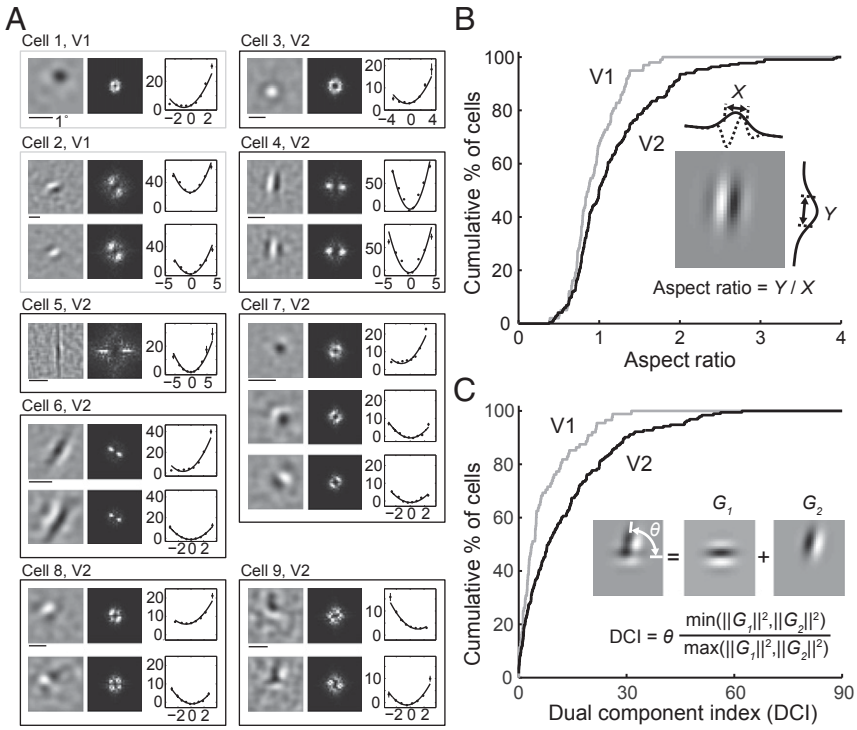

Fig. 2. RF subunits identified by PPR. (*A*) Example V1 and V2 RFs. Each box contains subunits from one cell (gray, V1; black, V2). Cell 1, V1 cell with a center-surround subunit. Eccentricity (Ecc), 2.4°. (*Left*) Linear filter. (*Center*) RF spectrum. (*Right*) Nonlinear function (data points in spikes/s; error bars, SEM; curves, polynomial fits). Cell 2, V1 cell with two Gabor subunits. Ecc, 7.6°. Cell 3, V2 cell with a center-surround subunit. Ecc, 3.7°. Cell 4, V2 cell with two Gabor subunits. Ecc, 2.2°. Cells 5 and 6, V2 cells with ultralong subunits. Ecc, 2.2° and 2.1°. Cells 7–9, V2 cells with complex-shaped RF subunits. Ecc, 2°, 3°, and 3.8°. (*B*) Cumulative distributions of aspect ratio for V1 (n = 99) and V2 (n = 223) cells. (*Inset*) Gabor fit for subunit 1 of cell 4. *X* and *Y*, widths of Gaussian envelope (SD) orthogonal and parallel to the sine wave grating. AR of each cell was averaged from all excitatory subunits. (*C*) Cumulative distributions of DCI for V1 (n = 88) and V2 (n = 193) cells. (*Inset*) Fit of subunit 2 of cell 9 by sum of two Gabors (G_1, G_2). θ, angle between two Gabors; $\|G\|^2$, $\iint G^2 dxdy$. The DCI of each cell was averaged from all excitatory subunits.

(Fig. S4*C*), and the latter's contribution was further assessed by comparing the measured responses with the predicted responses before and after their removal. We found increased deviation of the predicted from measured responses by removal of the second (Fig. S4 *A*, arrow, and *B*, right arrow) or third (Fig. S4*B*, left arrow) subunit, especially when the images contained spatial features that matched the complex-shaped subunits (Fig. S4 *A* and *B*, *Insets*). Removal of one or two high DCI subunits for high DCI cells (DCI > 30) also showed significant decrease of CC (Fig. S4*D*; one, $P = 0.004$; two, $P = 0.00002$, Wilcoxon signed-rank test).

Furthermore, we tested whether each oriented component in the complex-shaped RF subunits was important. As shown in Fig. S4*E*, removal of an oriented component from one subunit or from all subunits of the example cell reduced the model's performance measured by CC. Performing this analysis for all high DCI cells (DCI > 30) also showed a significant decrease of CC (Fig. S4*F*; $P_1 = 2 \times 10^{-6}$, $P_{all} = 9 \times 10^{-6}$, Wilcoxon signed-rank test). Thus, the non–V1-like RF structures identified with PPR contribute significantly to V2 neuronal representation of natural images.

Organization of RFs in V2 Functional Domains. A hallmark of neocortical circuits is the nonrandom spatial organization of neurons based on their functional properties. Previous studies on monkeys using cytochrome oxidase (CO) staining have delineated area V2 into alternating thin, pale, and thick stripes (25–27), in which color, orientation, and direction selective cells are preferentially located, respectively (1, 28–31). We next inquired whether cells with identified RF subunit structures are organized into different V2 stripes.

To target specific stripes for single-unit recording, we implanted a cranial chamber with penetrable artificial dura near the lunate sulcus (Fig. 3*A*), which allowed optical imaging of intrinsic signals and subsequent electrophysiological recordings from both V1 and V2 over a duration of 3–4 months. We measured four functional maps (color, orientation, direction, and ocular dominance) using intrinsic optical imaging under anesthesia (Fig. 3*B* and Fig. S5 *C*–*F*) and identified the thin (30), pale, and thick (31) stripes (Fig. 3*B* and *Materials and Methods*) that were previous identified on the basis of CO staining. The locations of 155 recorded V2 cells from three hemispheres of two monkeys were registered to the three types of stripes using blood vessels as landmarks (*Materials and Methods*).

Using the same RF analysis procedures described above, we found that cells with ultralong RFs (AR > 2; Fig. 4*A*; cells 3 and 4) were primarily located in the pale stripes (Fig. 4*B*), although both pale and thick stripes contained more high AR cells than V1 as shown by the cumulative distribution of ARs (Fig. 4*C*; V1 vs. V2 pale stripes: $P = 0.00002$; V1 vs. V2 thick stripes: $P = 0.03$; V2 thin and pale stripes: $P = 0.0001$; V2 thin and thick stripes: $P = 0.02$; Kolmogorov–Smirnov test). On the other hand, cells with complex-shaped RFs (high DCIs, Fig. 4*A*; cells 1 and 2) were located preferentially in thin stripes, as shown by cumulative distribution of DCIs (Fig. 4*D*), and a significant difference was found between V1 and V2 thin stripes ($P = 0.004$, Kolmogorov–Smirnov test). Simple cells with center-surround RFs were also concentrated in thin stripes (Fig. S6). Most cells in thick stripes had V1-like RFs (Fig. 4*A*; cells 5 and 6), which were widely distributed in all stripes.

Discussion

Using natural stimuli and the PPR method, we found that two types of RFs are prominent in V2 of awake macaque monkeys. The ultralong RFs have eluded detection using traditional visual stimuli. The existence of such cells in the visual system has been suggested by the human psychophysical evidence that visual contrast sensitivity could be enhanced by elongating but not widening of the Gabor stimulus (32). The neurons with complex-shaped RFs presumably correspond to previously identified RFs showing preference for multiple orientations (3–5) and RFs with 2D translation invariance to spot stimulus, known as complex-unoriented (23) or spot cells (24).

The function of visual neurons could be inferred by their RF properties. Intuitively, the ultralong cells integrate stimuli over a larger area along the orientation axis and thus could facilitate the perception of illusory contours by V2 (6). From the perspective of information processing, V1 cells reduce redundancy in the coding of natural scenes by maximizing the independence of outputs across neurons (33). Independent component analysis performed on the modeled outputs of V1 to natural stimuli resulted in V2 RFs with optimal stimulus more elongated than V1 cells (34), suggesting ultralong cells found here in V2 may further improve the coding efficiency for natural images.

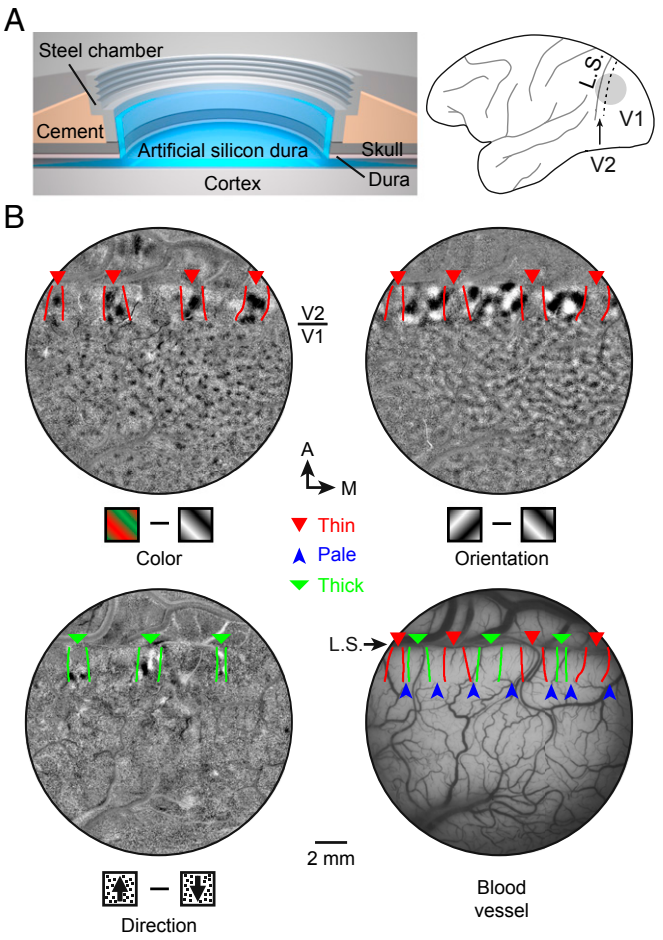

Fig. 3. Functional domains identified by optical imaging of intrinsic signals. (*A*) (*Left*) Cranial window for intrinsic optical imaging and single-unit recording. (*Right*) Gray circle, location of the window; dashed line, V1/V2 border; L.S., lunate sulcus. (*B*) Three functional maps (color, orientation, and direction) obtained by optical imaging of intrinsic signals and blood vessel pattern of the same cortical area, overlaid with expected stripe borders defined previously by cytochrome oxidase staining (*Materials and Methods*). Arrowheads, different stripes. A, anterior; M, medial.

The complex-shaped cells may be well suited for the detection of corners or T-junctions and thus providing information for surface analysis (35) and object recognition. A previous neural network modeling study identified V2 RFs with structures similar to complex-shaped cells found here and suggested these cells provide sparse coding for natural images (36). These cells may also provide translation-invariant representation for natural images, because independent component analysis on translation-invariant representation of natural image patches resulted in double-Gabor filters (37) that are reminiscent of some of the complex-shaped cells found in this study.

In addition to the ultralong and complex-shaped RFs in V2, we also observed many V2 RFs similar to those found in LGN and V1, consistent with previous electrophysiological studies (1, 2). These V2 cells had larger RF sizes but conserved the spatial shapes that could be well represented by a wavelet function [Laplacian of Gaussian for center-surround RF (38), and Gabor for V1 RF (20)], thus functioning as part of the wavelet representation of the stimulus at a coarser scale. This suggests that RFs of such simple properties might also exist in higher cortical areas with larger RF sizes, such as V4 and IT.

As well as the RF properties described by the subunit structures, V2 neurons are also likely to exhibit other nonlinear properties that cannot be captured by the simple feedforward model, such as discrimination of border ownership (9), which may also rely on feedback signals from higher cortical areas.

To locate the V2 stripes, we used optical imaging of intrinsic signal in vivo. Further postmortem identification with CO staining was not performed because the monkeys are still alive. The colocalization of CO thin stripes and color domains of intrinsic signal was previously verified by CO staining (30), whereas other evidence from intrinsic signals suggests that direction domains fall within thick stripes (31). Based on these previous findings, our determination of thin stripes is likely to be accurate, but that of thick stripes may be an underestimate. Strictly speaking, our results reflect the preferential localization of cells with respect to functional stripes defined by intrinsic imaging rather than anatomical (CO) staining. Furthermore, the finding that different RFs are preferentially distributed in different stripes fits the idea that thin and pale stripes in V2 are part of the ventral pathway important for form processing and object recognition (39–45).

As exemplified by the findings in the early visual pathway, neurons in the retina and LGN mostly have RFs with center-surround ON/OFF regions (46), and V1 cells could integrate linearly aligned RFs of adjacent LGN neurons to form RFs with alternating on/off regions with a preferred orientation (47). Thus, characterization of the spatial structure of RFs in higher cortical areas will allow us to understand how input signals become integrated progressively through multiple stages to eventually achieve visual functions such as object recognition. Because the feedforward input of V2 originates primarily from V1 (48–51), we could infer the rule of the convergence from V1 to V2 cells from the RF properties of V2 cells. The finding that V2 RF subunits could be well fitted by the summation of multiple Gabor functions (Fig. S2) supports a simple

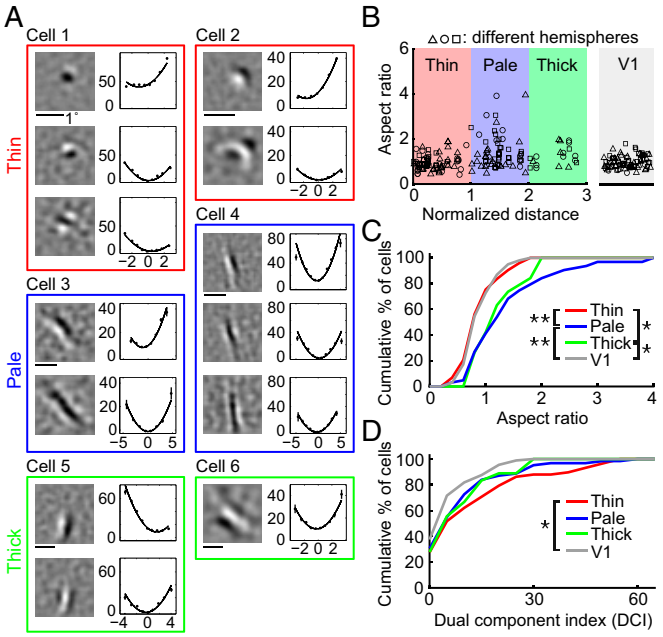

Fig. 4. Organization of RF properties in V2 stripes. (*A*) Example RFs of six cells from thin (red), pale (blue), and thick (green) stripes. Eccentricity, 1.9°, 1.7°, 2.0°, 1.7°, 1.9°, and 0.9°. (*B*) (*Left*) Aspect ratio vs. normalized location across stripes (0, 1.5, and 3 represent center of thin, pale, and thick stripes, respectively). Data point with different symbol, aspect ratio of neuron from different hemispheres. (*Right*) Aspect ratios of V1 cells (randomly dispersed horizontally). (*C*) Cumulative distribution of aspect ratios of V1 ($n = 99$) and V2 cells in thin ($n = 73$), pale ($n = 63$), and thick ($n = 19$) stripes. Significant differences were found between V1 and V2 cells in pale or thick stripes and between V2 cells in thin and pale or thick stripes. (*D*) Cumulative distributions of DCI of V1 ($n = 88$) and V2 cells in thin ($n = 58$), pale ($n = 62$), and thick ($n = 18$) stripes. Significant difference was found between V1 and V2 cells in thin stripes.

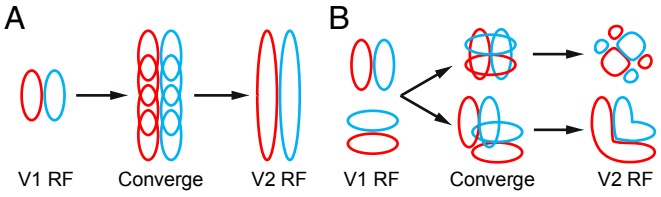

Fig. 5. A convergence model for V2 RFs. (*A*) The model for V2 cells with ultralong RF subunits. Red and blue ellipse, on and off subregions. An ultralong cell may receive convergent projections from multiple V1 cells aligned along their preferred orientation. (*B*) The model for V2 cells with complex-shaped RF subunits. The cells may receive convergent inputs from V1 cells with different orientations at same or different center locations.

convergence model (Fig. 5), in which each V2 neuron receives convergent inputs from several V1 cells with similar or different orientations, allowing V2 neurons to detect elongated edges, corners, and complex textures in natural scenes.

Materials and Methods

Surgical Preparation. Three adult male macaque monkeys (*Macaca mulatta*) were used for this study. All procedures were in accordance with National Institutes of Health guidelines and approved by the Institutional Animal Care and Use Committee of Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences. Surgery for chronic chamber installation (at lunate sulcus, ~16 mm anterior to occipital bone ridge and ~13 mm lateral from midline) were performed using standard sterile techniques under deep anesthesia maintained by artificial respiration of isoflurane (1–2.5%). For combined electrical and optical recording in monkeys B and C, we implanted transparent electrode-penetrable silicone "artificial dura" (inner diameter, 18 mm; Fig. 4*A*) (52, 53), which protected and stabilized the cortex and prevented dura regrowth for a period up to 3–4 months.

Electrophysiological Recording. Single-unit recordings were made in both V1 and V2 of four hemispheres in three awake monkeys using glass-coated tungsten electrodes while the monkeys were doing a fixation task for a juice reward. Eye position was monitored with a remote infrared eye tracker (EYELINK II; sampling rate, 500 Hz; SR Research). Data recorded when the eye position was outside of an elliptical window (size: vertical 1.2°, horizontal 1°) were excluded. All single units presented with >12,000 images during fixation were included in the analysis (124 V1 and 360 V2 cells). The eccentricity of RFs ranges from 1° to 8°. Single units were isolated using cluster analysis of spike waveforms (Offline Sorter; Plexon) and confirmed by the refractory period in the histogram of interspike intervals.

Visual stimuli were generated with a windows-based PC, presented on a gamma-calibrated CRT monitor (Sony Multiscan G520; 30 × 40 cm, refresh rate, 100 Hz; maximum luminance, 80 cd/m^2). Classical RF of each neuron was estimated using a drifting sinusoidal grating, with its spatial and temporal frequencies chosen to effectively drive the cell. Natural images (16 × 16 or 32 × 32 pixels) sampled randomly from the database of digitized movies (19) were displayed at a frame rate of 20 Hz. The mean luminance and global contrast were normalized for all selected images.

Cell Localization. For monkey A, the identification of V1 and V2 areas was based on their retinotopic maps, and cells close to the V1/V2 border (with RF near the middle line of visual field) were excluded. The recordings from monkeys B and C were guided by optical imaging of intrinsic signals, and the precise location of the recording site was determined by injecting a small amount of dye (fast green) into the guide tube after each recording to mark the electrode location on the artificial dura. The mark was registered to the stripes identified from intrinsic imaging map, based on prominent blood vessel branch points.

Data Partitioning. The recorded responses of each cell were randomly partitioned into training data (80%, for determining RF model) and test data (20%, for evaluating RF model obtained). Training data were further partitioned into four segments with alternating 3/4 for computing model parameters and 1/4 as validation data to determine the final RF model from multiple models derived by PPR (see below).

Linear–Nonlinear Model. The linear–nonlinear model (14, 15, 54) used to represent the RF is

$$R = \sum_{i=1}^{n} f_i(S \cdot F_i) \quad [1]$$

where R represents the stimulus-evoked firing rate of the neuron, S is the stimulus, represented by the luminance of pixels, F_i is the spatial filter of the ith subunit (with same dimension as S), and f_i represents the static nonlinear function $f_i(x_i) = a_i + b_i\, x_i + c_i\, x_i^2$, with a, b, and c as free parameters. The parameters of the model for each cell were determined by the following least square optimization (17)

$$\min \sum (r - R)^2 + \lambda g(F), \quad [2]$$

where r is the measured neuronal response, and R is the response predicted by the model (Eq. **1**). To avoid overfitting, the second term was used to penalize the nonsmoothness of the spatial filter F, as measured by $g(F)$ as follows:

$$g(F) = \sum_i \left\| \nabla^2 F_i \right\|^2, \quad [3]$$

where ∇^2 represents a 3 × 3 Laplacian operator, and λ is a scalar parameter.

PPR. The PPR method (16–18) used two steps to determine the RF model. In the forward step, it searched for one subunit at a time. The search started with a random filter as the first subunit by fitting the output of the subunit to the recorded response using gradient descent. Following identification of each subunit, it searched for the next subunit by fitting the residual neural responses (measured responses minus the output of all existing subunits). When the number of subunit reached the designated N (20 used in this study), the forward step was terminated. In the backward step, PPR simplified and refined the RF model. First, the subunits were sorted by their contribution to the model evaluated by their output variance. The subunits were then eliminated one at a time (starting by the one with least contribution), and the remaining subunits were refined after each elimination by fitting them with the residual neuronal response. The backward step was terminated when only one subunit was left. For each PPR computation we obtained $N - 1$ models from the backward step, with each model having 1 to $N - 1$ subunits. For each cell, we computed PPR at various λ values and two different time delays. Each model was then evaluated by CC between predicted and measured responses of 20 segments of validation data (random segmented with equal length). The final model was chosen as the one with the highest average CC, among models with the least number of subunits and CCs not significantly lower than the model with highest CCs of all models ($P > 0.01$, Wilcoxon signed-rank test).

Quantitative Analysis of RF Structures. Before the analysis of the spatial RF structure, the final model for each cell was evaluated by CC between the predicted and measured responses of test data. The models with CC < 0.1 were excluded from the analysis. Further characterization of the RF subunits was done by evaluating the least square fit of all of the spatial filters of excitatory subunits with the following functions: DOG (38), 2DGabor, and sum of two 2D Gabors. The goodness-of-fit for DOG fitting was calculated as $||F||^2/(||DOG - F||^2)$ within the RF region (radius = 4 SD of the fitted DOG function), where F is the spatial RF of the cell.

Comparison of PPR Method with Other Methods. We compared the performance of PPR method with other computational methods in characterizing V2 RFs. Our initial attempt of using maximally informative dimension (MID) method (55) yielded unsatisfactory results for RFs with multiple subunits, although its performance was good for those with a single subunit. We then turned to the modified spike-triggered average or spike-triggered covariance (mSTA/STC) method for comparison with PPR. The mSTA/STC analysis for natural stimuli have been described previously (21, 56–58). For comparison, we constructed the RF models by mSTA/STC with the same number of subunits as those from PPR for each cell. First, the filters were computed at various parameters of regularization α and two different time delays using mSTA/STC. To evaluate the importance of each filter, single-subunit models were constructed from these filters (the second-order polynomial function was estimated for each filter using training data). Then, for each α and time delay of a cell, we selected the best n filters, whose single-subunit model predicted better than others, to construct an n-subunit model, with n equal to the number of subunits of RF model resulting from PPR. The nonlinear functions were estimated again for the n-subunit model by linear regression on training data. Finally, the model for each cell was chosen as the one with the best prediction of validation data.

Optical Imaging. The standard method (59) of optical imaging of intrinsic signals for anesthetized animal was performed on three hemispheres of two

monkeys (B and C) in an 20 × 16.5- or 17 × 14-mm area of the cortex, with 654 × 540-pixel resolution. Visual stimuli were repeated for 30–50 blocks, and each block contained several sets of stimulus conditions (e.g., gratings with different orientations) and a blank screen of the same mean luminance (8 cd/m^2, 8 s). Visual stimuli for optical imaging were created with ViSaGe (Cambridge Research Systems) and presented on a gamma-corrected CRT monitor (SONY CPD-G520). Full screen drifting gratings (spatial frequency, 1.5 cycle/°; temporal frequency, 8 Hz) were presented in a randomly interleaved manner in eight directions (0°, 45°,...315°) to obtain orientation and ocular dominance maps. Electromechanical shutters were used for monocular stimulation to measure ocular dominance maps. To obtain a color preference map, the stimuli comprised of red-green isoluminance sine-wave gratings and 100% contrast black-and-white gratings were presented in one of two orientations (45°, 135°) at random directions. Red Commision Internationale de L'Eclairage (CIE) values were 0.552 and 0.299; green CIE values were 0.268 and 0.530. The black-white gratings were set to have the same spatial frequency, temporal frequency, and average screen luminance as red-green gratings. Moving random white dots (0.04°) drifted at 8°/s in one of eight directions (0°, 45°,...315°) with a density of 10% of the monitor area were used to obtain direction maps.

For each stimulus condition, we first obtained single-condition maps, which represent percent changes in intrinsic signals relative to the prestimulus baseline as $dR/R = (R_i - R_{1–3})/R_{1–3}$, where $R_{1–3}$ is the average raw reflectance value of frames 1–3 (taken before stimulus onset and representative of the baseline activity), and R_i represents a single frame between frame 8 and 16. Then, these single-condition maps were used for calculating support vector machine (SVM) maps (60) (Fig. 3*B* and Fig. S5 *C*–*F*), which have higher signal-to-noise ratio than traditional subtraction maps. These SVM maps were band-pass filtered and normalized for determining the V2 stripe borders, defined as the contours of the cross-points between color/orientation and orientation/direction maps (Fig. S5*G*).

ACKNOWLEDGMENTS. We thank Dr. Yunqing Wen, Dr. Hou Xu, Dr. Xiaodong Chen, Junjie Cai, Dr. Shude Zhu, Dr. Peichao Li, and Dr. Yu-Xi Fu for technical support and Dr. Haishan Yao and Dr. Wei Wang for comments. This work was supported by grants from Ministry of Science and Technology (973 Program, 2011CBA00400) and Chinese Academy of Sciences (Strategic Priority Research Program, XDB02020001).

1. Hubel DH, Livingstone MS (1987) Segregation of form, color, and stereopsis in primate area 18. *J Neurosci* 7(11):3378–3415.
2. Levitt JB, Kiper DC, Movshon JA (1994) Receptive fields and functional architecture of macaque V2. *J Neurophysiol* 71(6):2517–2542.
3. Ito M, Komatsu H (2004) Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *J Neurosci* 24(13):3313–3324.
4. Anzai A, Peng X, Van Essen DC (2007) Neurons in monkey visual area V2 encode combinations of orientations. *Nat Neurosci* 10(10):1313–1321.
5. Tao X, et al. (2012) Local sensitivity to stimulus orientation and spatial frequency within the receptive fields of neurons in visual area 2 of macaque monkeys. *J Neurophysiol* 107(4):1094–1110.
6. von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. *Science* 224(4654):1260–1262.
7. Hegdé J, Van Essen DC (2000) Selectivity for complex shapes in primate visual area V2. *J Neurosci* 20(5):RC61.
8. Freeman J, Ziemba CM, Heeger DJ, Simoncelli EP, Movshon JA (2013) A functional and perceptual signature of the second visual area in primates. *Nat Neurosci* 16(7):974–981.
9. Zhou H, Friedman HS, von der Heydt R (2000) Coding of border ownership in monkey visual cortex. *J Neurosci* 20(17):6594–6611.
10. Marcus DS, Van Essen DC (2002) Scene segmentation and attention in primate cortical areas V1 and V2. *J Neurophysiol* 88(5):2648–2658.
11. Sharpee TO (2013) Computational identification of receptive fields. *Annu Rev Neurosci* 36:103–120.
12. Wu MC-K, David SV, Gallant JL (2006) Complete functional characterization of sensory neurons by system identification. *Annu Rev Neurosci* 29:477–505.
13. Felsen G, Touryan J, Han F, Dan Y (2005) Cortical sensitivity to visual features in natural scenes. *PLoS Biol* 3(10):e342.
14. de Boer R, Kuyper P (1968) Triggered correlation. *IEEE Trans Biomed Eng* 15(3):169–179.
15. Carandini M, et al. (2005) Do we know what the early visual system does? *J Neurosci* 25(46):10577–10597.
16. Friedman JH, Stuetzle W (1981) Projection pursuit regression. *J Am Stat Assoc* 76(376): 817–823.
17. Rapela J, Felsen G, Touryan J, Mendel JM, Grzywacz NM (2010) ePPR: A new strategy for the characterization of sensory cells from input/output data. *Network* 21(1-2):35–90.
18. Rapela J, Mendel JM, Grzywacz NM (2006) Estimating nonlinear receptive fields from natural images. *J Vis* 6(4):441–474.
19. van Hateren JH, van der Schaaf A (1998) Independent component filters of natural images compared with simple cells in primary visual cortex. *Proc Biol Sci* 265(1394):359–366.
20. Jones JP, Palmer LA (1987) An evaluation of the two-dimensional Gabor filter model of simple receptive fields in cat striate cortex. *J Neurophysiol* 58(6):1233–1258.
21. Touryan J, Felsen G, Dan Y (2005) Spatial structure of complex cell receptive fields measured with natural images. *Neuron* 45(5):781–791.
22. Chen X, Han F, Poo M-M, Dan Y (2007) Excitatory and suppressive receptive field subunits in awake monkey primary visual cortex (V1). *Proc Natl Acad Sci USA* 104(48):19120–19125.
23. Hubel DH, Livingstone MS (1985) Complex-unoriented cells in a subregion of primate area 18. *Nature* 315(6017):325–327.
24. Baizer JS, Robinson DL, Dow BM (1977) Visual responses of area 18 neurons in awake, behaving monkey. *J Neurophysiol* 40(5):1024–1037.
25. Livingstone MS, Hubel DH (1982) Thalamic inputs to cytochrome oxidase-rich regions in monkey visual cortex. *Proc Natl Acad Sci USA* 79(19):6098–6101.
26. Tootell RB, Silverman MS, De Valois RL, Jacobs GH (1983) Functional organization of the second cortical visual area in primates. *Science* 220(4598):737–739.
27. Horton JC (1984) Cytochrome oxidase patches: A new cytoarchitectonic feature of monkey visual cortex. *Philos Trans R Soc Lond B Biol Sci* 304(1119):199–253.
28. Gegenfurtner KR, Kiper DC, Fenstemaker SB (1996) Processing of color, form, and motion in macaque area V2. *Vis Neurosci* 13(1):161–172.
29. Shipp S, Zeki S (2002) The functional organization of area V2, I: Specialization across stripes and layers. *Vis Neurosci* 19(2):187–210.
30. Lu HD, Roe AW (2008) Functional organization of color domains in V1 and V2 of macaque monkey revealed by optical imaging. *Cereb Cortex* 18(3):516–533.
31. Lu HD, Chen G, Tanigawa H, Roe AW (2010) A motion direction map in macaque V2. *Neuron* 68(5):1002–1013.
32. Polat U, Tyler CW (1999) What pattern the eye sees best. *Vision Res* 39(5):887–895.
33. Bell AJ, Sejnowski TJ (1997) The "independent components" of natural scenes are edge filters. *Vision Res* 37(23):3327–3338.
34. Hyvärinen A, Gutmann M, Hoyer PO (2005) Statistical model of natural stimuli predicts edge-like pooling of spatial frequency channels in V2. *BMC Neurosci* 6:12.
35. Rubin N (2001) The role of junctions in surface completion and contour matching. *Perception* 30(3):339–366.
36. Lee H, Ekanadham C, Ng AY (2008) Sparse deep belief net model for visual area V2, eds Platt JC, Koller D, Singer Y, Roweis ST, *Advances in Neural Information Processing Systems* (Curran Associates, Red Hook, NY), Vol 20, pp 873–880.
37. Saremi S, Sejnowski TJ, Sharpee TO (2013) Double-gabor filters are independent components of small translation-invariant image patches. *Neural Comput* 25(4):922–939.
38. Rodieck RW (1965) Quantitative analysis of cat retinal ganglion cell response to visual stimuli. *Vision Res* 5(11):583–601.
39. Ungerleider LG, Mishkin M (1982) Two cortical visual systems. *Analysis of Visual Behavior*, eds Ingle DJ, Goodale MA, Mansfield RJW (MIT Press, Cambridge, MA), pp 549–586.
40. DeYoe EA, Felleman DJ, Van Essen DC, McClendon E (1994) Multiple processing streams in occipitotemporal visual cortex. *Nature* 371(6493):151–154.
41. Connor CE, Brincat SL, Pasupathy A (2007) Transformation of shape information in the ventral pathway. *Curr Opin Neurobiol* 17(2):140–147.
42. Logothetis NK, Sheinberg DL (1996) Visual object recognition. *Annu Rev Neurosci* 19: 577–621.
43. Tsao DY, Livingstone MS (2008) Mechanisms of face perception. *Annu Rev Neurosci* 31:411–437.
44. DiCarlo JJ, Zoccolan D, Rust NC (2012) How does the brain solve visual object recognition? *Neuron* 73(3):415–434.
45. DeYoe EA, Van Essen DC (1985) Segregation of efferent connections and receptive field properties in visual area V2 of the macaque. *Nature* 317(6032):58–61.
46. Kuffler SW (1953) Discharge patterns and functional organization of mammalian retina. *J Neurophysiol* 16(1):37–68.
47. Hubel DH, Wiesel TN (1962) Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol* 160:106–154.
48. Kuypers HGJM, Szwarcbart MK, Mishkin M, Rosvold HE (1965) Occipitotemporal corticocortical connections in the rhesus monkey. *Exp Neurol* 11(2):245–262.
49. Van Essen DC, Newsome WT, Maunsell JH, Bixby JL (1986) The projections from striate cortex (V1) to areas V2 and V3 in the macaque monkey: Asymmetries, areal boundaries, and patchy connections. *J Comp Neurol* 244(4):451–480.
50. Schiller PH, Malpeli JG (1977) The effect of striate cortex cooling on area 18 cells in the monkey. *Brain Res* 126(2):366–369.
51. Girard P, Bullier J (1989) Visual activity in area V2 during reversible inactivation of area 17 in the macaque monkey. *J Neurophysiol* 62(6):1287–1302.
52. Chen LM, et al. (2002) A chamber and artificial dura method for long-term optical imaging in the monkey. *J Neurosci Methods* 113(1):41–49.
53. Arieli A, Grinvald A, Slovin H (2002) Dural substitute for long-term imaging of cortical activity in behaving monkeys and its clinical implications. *J Neurosci Methods* 114(2):119–133.
54. Marmarelis VZ (2004) *Nonlinear Dynamic Modeling of Physiological Systems* (John Wiley & Sons, Inc., Hoboken, NJ).
55. Sharpee T, Rust NC, Bialek W (2004) Analyzing neural responses to natural signals: Maximally informative dimensions. *Neural Comput* 16(2):223–250.
56. Smyth D, Willmore B, Baker GE, Thompson ID, Tolhurst DJ (2003) The receptive-field organization of simple cells in primary visual cortex of ferrets under natural scene stimulation. *J Neurosci* 23(11):4746–4759.
57. Woolley SMN, Gill PR, Theunissen FE (2006) Stimulus-dependent auditory tuning results in synchronous population coding of vocalizations in the songbird midbrain. *J Neurosci* 26(9):2499–2512.
58. Sharpee TO, Miller KD, Stryker MP (2008) On the importance of static nonlinearity in estimating spatiotemporal neural filters with natural stimuli. *J Neurophysiol* 99(5):2496–2509.
59. Li P, et al. (2013) A motion direction preference map in monkey V4. *Neuron* 78(2): 376–388.
60. Xiao Y, Rao R, Cecchi G, Kaplan E (2008) Improved mapping of information distribution across the cortical surface with the support vector machine. *Neural Netw* 21(2-3):341–348.

Supporting Information

Liu et al. 10.1073/pnas.1525505113

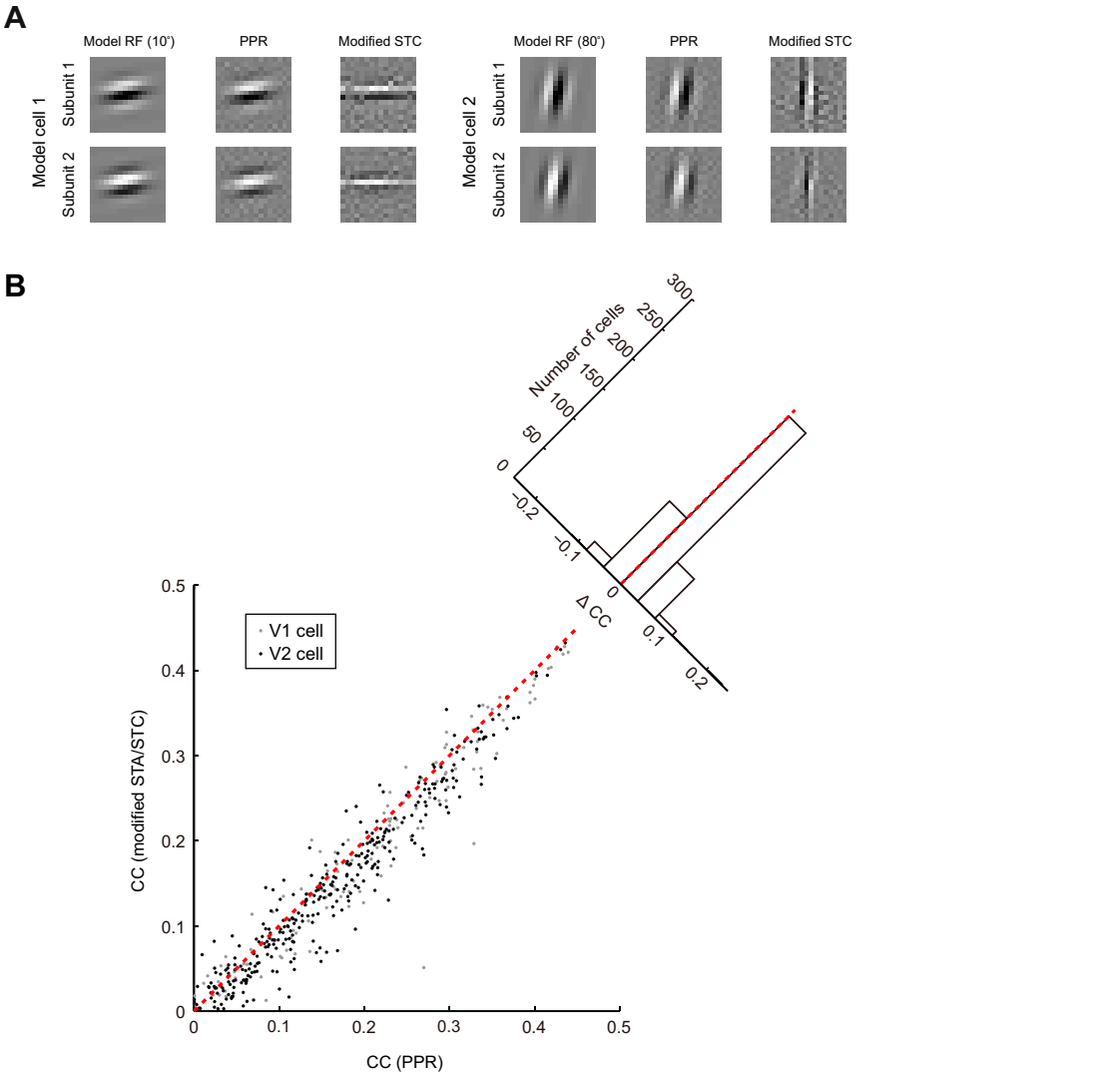

Fig. S1. Comparison of PPR with modified spike-triggered average or spike-triggered covariance (mSTA/STC) method. (*A*) For both model cells: (*Left*) RFs of model cells; (*Center*) RFs estimated by PPR; (*Right*) RFs estimated by modified STC. The distortion of RF by modified STC was most prominent when the orientation of model cell was near horizontal or vertical (e.g., orientation 10° or 80°). (*B*) Comparison of CCs for measured (test data) and predicted responses between RFs computed by mSTA/STC and PPR for both V1 (gray, $n = 124$) and V2 (black, $n = 360$) cells. Each dot represents one cell. Red dashed line, unit line. (*Inset*) Distribution of delta CC between PPR and mSTA/STC.

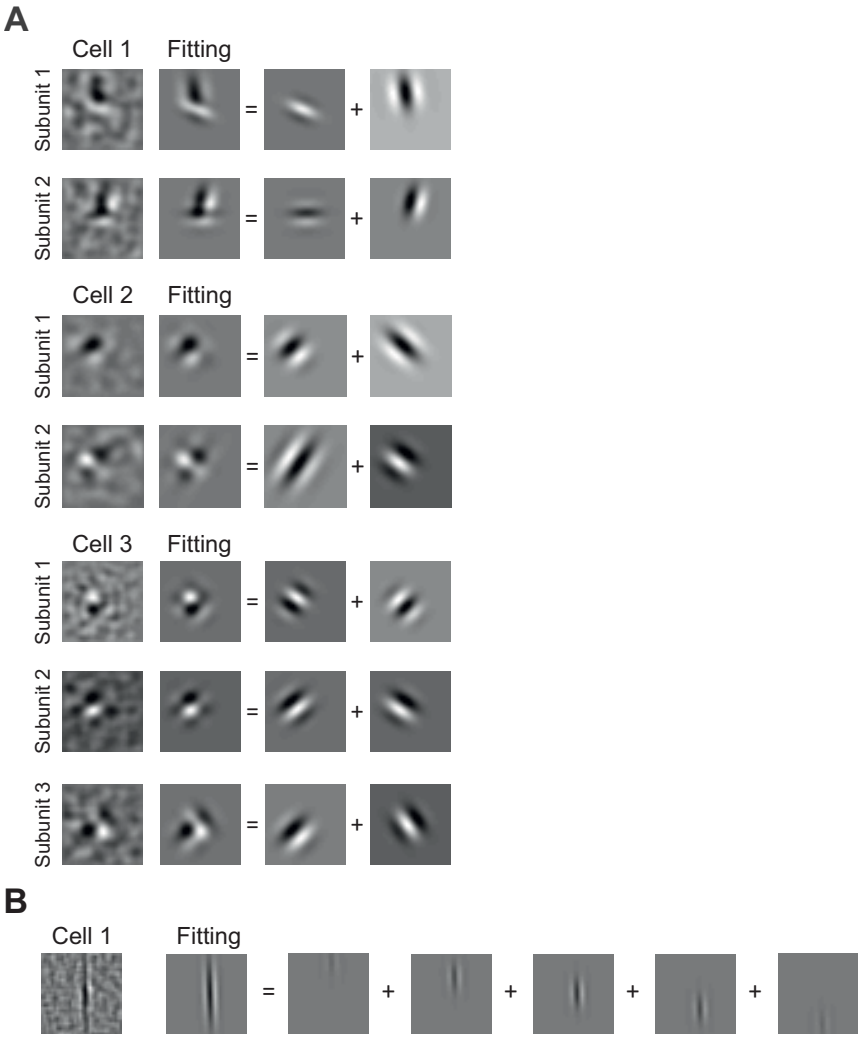

Fig. S2. Illustration of decomposing V2 complex-shaped RFs and ultralong RFs into sum of Gabor functions. (*A*) Three example V2 cells with complex-shaped subunits, fitted by the sum of two Gabor functions. (*B*) An example V2 cell with ultralong RF, fitted by sum of five equal spaced Gabor functions with same size.

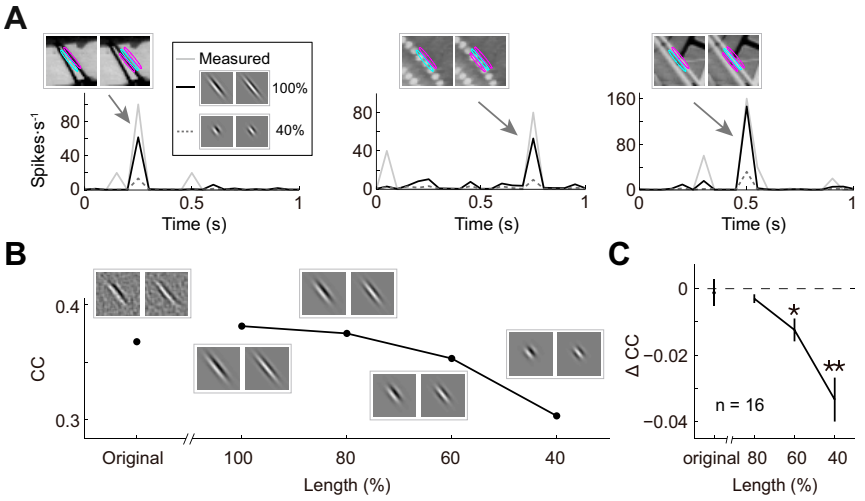

Fig. S3. Performance of RF models with ultralong subunits. (*A*) Prediction of responses by Gabor-fits of an ultralong RF (two subunits) and by shortened Gabors (40% of original). Arrows, example responses to the above stimuli, which was showed overlapping with RF contours. Note that all stimuli contained long edges aligned with the RF. (*B*) CCs between measured and predicted responses (of test dataset) by original RF subunits (computed from PPR), Gabor fits of the subunits, and progressively shortened Gabor. (*C*) Summary of analysis in *B* for all cells with AR > 2 ($n = 16$) by comparing the CC change with the best Gabor fits. The CC decrease is significant by shortening to 60% and 40%.

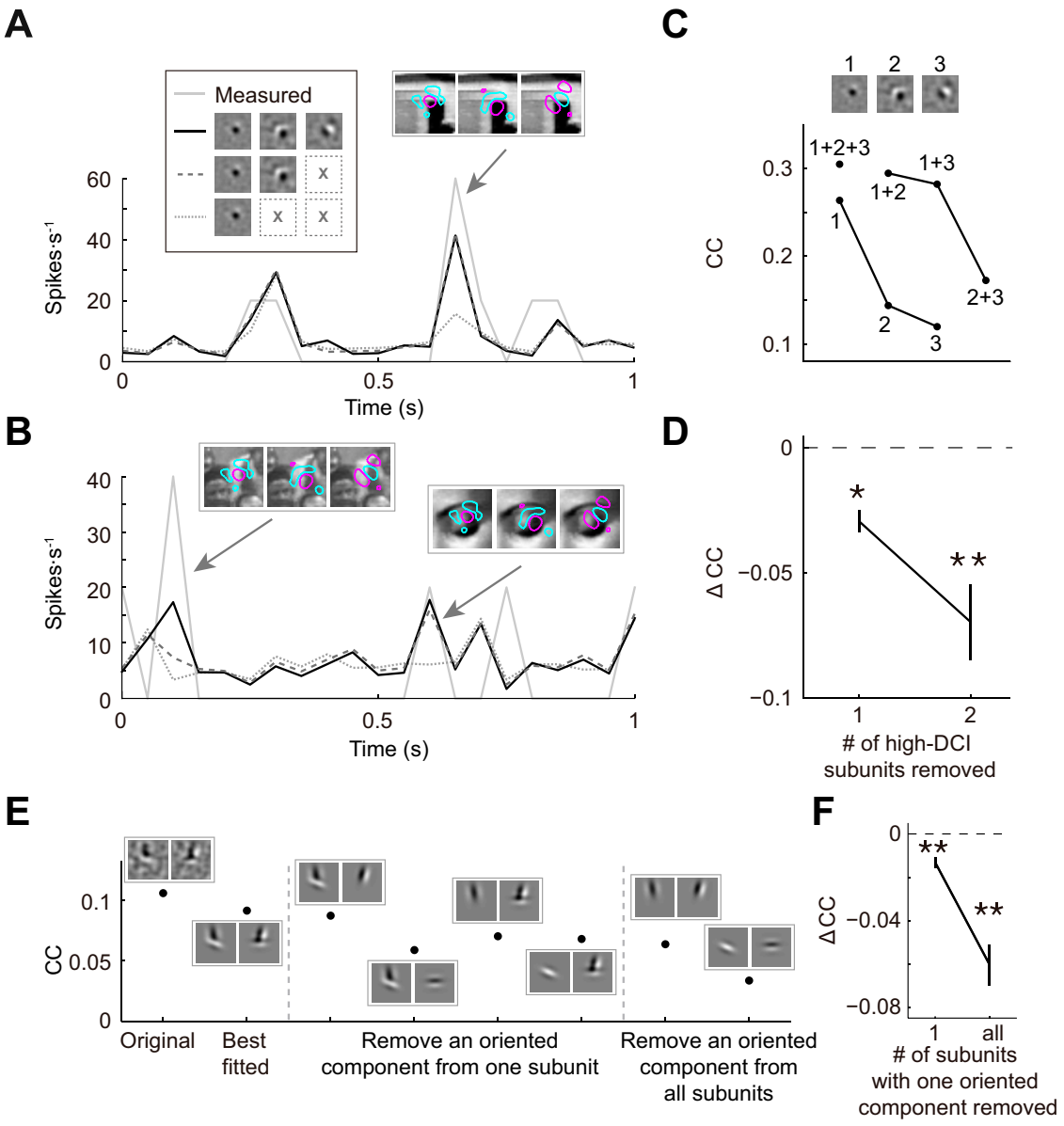

Fig. S4. Performance of RF models with complex-shaped subunits. (*A*–*D*) Prediction of responses before and after removal of subunits from cells with complex-shaped RF. (*A* and *B*) The example stimuli contained image features matched to one of the three subunits in the example V2 cell. Arrows, example responses to the above stimuli, which was showed overlapping with RF contours. (*C*) Three subunits of an example V2 cell are marked 1–3. Removal of the subunits reduced the CCs between measured (test data) and predicted responses. (*D*) Removal of one ($n = 24$) or two ($n = 9$) high DCI subunits from all high DCI cells (DCI > 30). (*E* and *F*) Prediction of responses before and after removing components of complex-shaped subunits. Removal of an oriented component from one subunit or from all subunits reduced the CCs between measured and predicted responses for an example cell (*E*) and all high DCI cells (*F*) (from one subunit, $n_1 = 48$; from all subunits, $n_{all} = 30$).

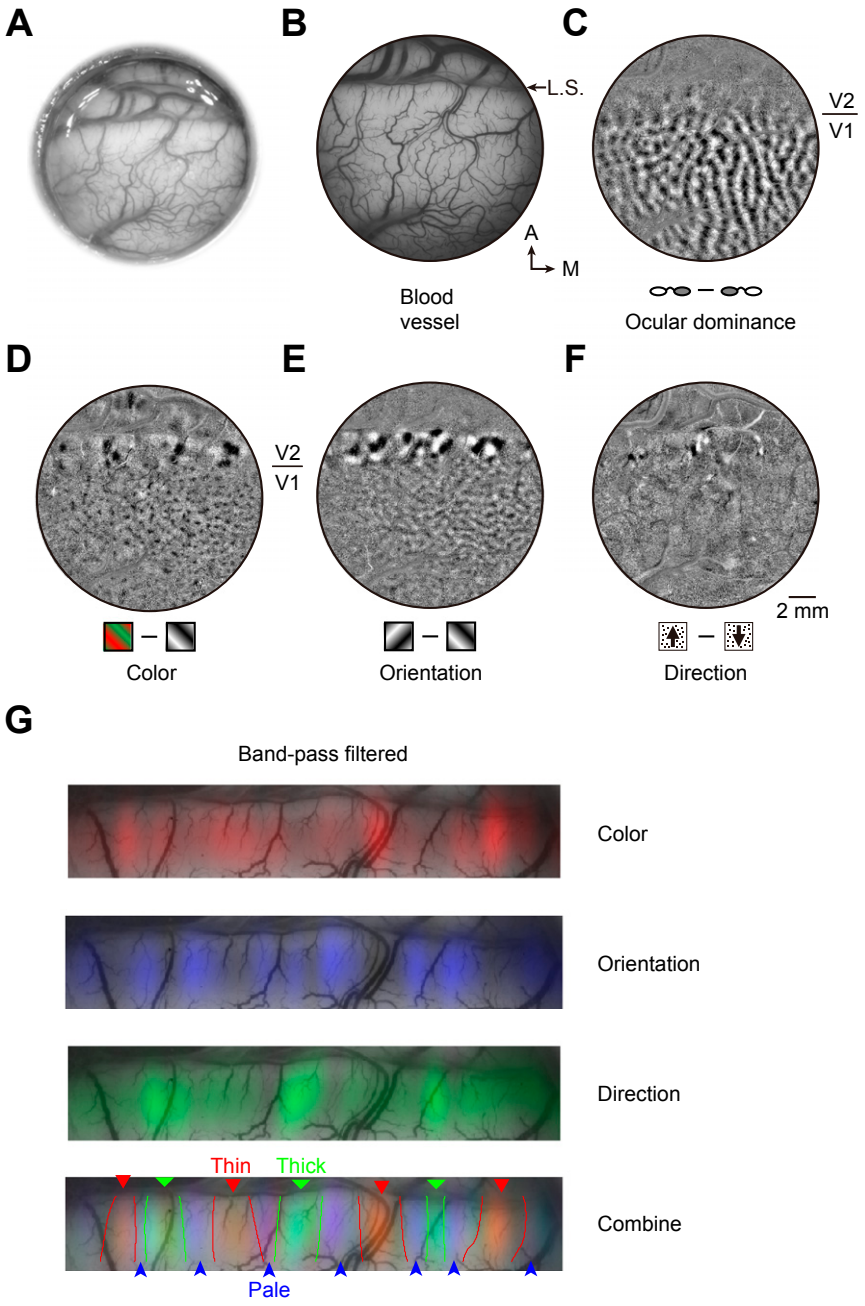

Fig. S5. Identification of expected stripes in V2 from functional maps obtained by optical imaging of intrinsic signals. (*A*) Bright-field image of the cortex through the implanted window. (*B*) Green light image of the visual cortex showing the pattern of blood vessels. A, anterior; M, medial; L.S., lunate sulcus. (*C*) Ocular dominance map, which is clear in V1 but not in V2. (*D*) Color map shows the blobs in V1 and color preference domains within thin stripes of V2. (*E*) Orientation map showing orientation-selective domains in both V1 and V2. In V2, these domains are found in thick and pale stripes, complementary to the color preference domains. (*F*) Motion direction map, indicating direction-selective domains in thick stripes of V2. (Scale bar, 2 mm for *B*–*F*.) (*G*) Band-pass filtered color, orientation, and direction maps overlaid with blood vessels in V2. (*Bottom*) Three overlaid maps, together with estimated stripe borders (*Materials and Methods*).

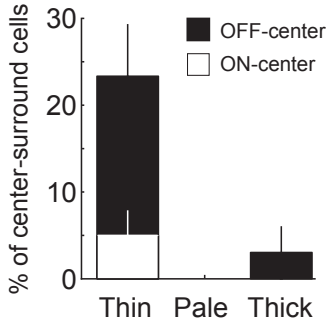

Fig. S6. Percentage of center-surround cells in different stripes. Most of these cells were located in thin stripes. Error bar, SEM (n = 3 hemispheres). Cells were defined as center-surround simple cell by two criteria: (*i*) containing only one subunit with half-wave-rectified nonlinear function; and (*ii*) the goodness-of-fit > 4 by fitting the RF with DOG function (*Materials and Methods*).