Current Biology, Vol. 14, 996–1001, June 8, 2004, ©2004 Elsevier Ltd. All rights reserved. DOI 10.1016/j.cub.2004.05.045

One Picture Is Worth at Least a Million Neurons

Ifat Levy,[1,2,3] Uri Hasson,[2] and Rafael Malach[2,*]
[1]The Interdisciplinary Center for
Neural Computation
Hebrew University of Jerusalem
Jerusalem 91904
Israel
[2]Department of Neurobiology
Weizmann Institute of Science
Rehovot 76100
Israel

Summary

How many neurons participate in the representation of a single visual image? Answering this question is critical for constraining biologically inspired models of object recognition, which vary greatly in their assumptions from few "grandmother cells" [1] to numerous neurons in widely distributed networks [2]. Functional imaging techniques, such as fMRI, provide an opportunity to explore this issue, since they allow the simultaneous detection of the entire neuronal population responding to each stimulus. Several studies [3–6] have shown that fMRI BOLD signal is approximately proportional to neuronal activity. However, since it provides an indirect measure of this activity, obtaining a realistic estimate of the number of activated neurons requires several intervening steps. Here, we used the extensive knowledge of primate V1 to yield a conservative estimate of the ratio between hemodynamic response and neuronal firing. This ratio was then used, in addition to several cautious assumptions, to assess the number of neurons responding to a single-object image in the entire visual cortex and particularly in object-related areas. Our results show that at least a million neurons in object-related cortex and about two hundred million neurons in the entire visual cortex are involved in the representation of a single-object image.

Results

To calculate a lower bound for the number of neurons that participate in the representation of a single image, we used well-established data regarding V1 and occipito-temporal cortex anatomy and physiology in combination with new experimental results (Figure 1).

Three stimuli were used in an event-related designed experiment. A moving rectangular-wave grating was used to maximally activate V1 neurons sensitive to a particular orientation [7]. Images of a single face and a single house were used to activate object-related areas. Subjects had to covertly name each stimulus. Despite the use of a single image, the resultant activation was quite large and widespread in occipital and temporal cortex (Figure 2A). The grating elicited substantial activation in V1, as well as in higher-order areas (Figure 2A, left). The face image typically preferentially activated a region in the posterior fusiform gyrus (pFs) corresponding to the FFA [8], whereas the house image activated a region in the collateral sulcus (CoS), sometimes extending into the parahippocampal gyrus, corresponding to the PPA [9] (Figure 2A, right). Both images also activated additional areas in occipito-temporal cortex as well as lower-tier areas [10].

Below we describe the analysis steps and the assumptions used in the estimation of activated neurons (see Figure 1).

Step One: V1 Neuronal Activity

Several parameters were used for this estimation. (1) The neuronal density of V1 was estimated, based on a recent study done in a large population of postmortem human brains [11], as about 60,000 neurons/mm^3. (2) The fraction of neurons, which may respond to the moving oriented rectangular grating by firing significantly above their baseline, was estimated based on electrophysiological studies [12, 13]. Some neuronal populations may not have been sensitive to oriented gratings, but to other stimulus attributes (e.g., specific colors [12]). Other populations may be tuned to spatial frequencies not contained in our stimulus [13]. Considering these facts, we conservatively estimated the fraction of responsive neurons to the grating as between 20%–50%. (3) The response to a grating of a single orientation, out of the maximal possible response generated by all possible orientations, was assessed based on the orientation tuning curve of V1 neurons. The average tuning curve half width at half maximum (HWHM) was estimated as between 27° (based on monkey single units [13, 14]) and ~45° (based on deoxyglucose uptake [15] optical imaging in anaesthetized macaque [16], and human fMRI [7]). The integral of the tuning curve yields the population "spike contribution" of a single orientation grating. Assuming a Gaussian tuning curve and a homogeneous distribution of orientation preferences, the average spike contribution is between 32% and 60% of the overall possible response (which is based on maximal possible firing of all responsive neurons; see Supplemental Experimental Procedures). Combined with the estimate of orientation-responsive neurons, this gives an overall percentage of between 6% and 30% of the maximal response of the entire population. (4) To obtain the maximal response of the entire population, the V1 peak firing rate was estimated, based on electrophysiological data, as about 45 spikes/s [17].

Multiplying V1 neuronal density by the percentage of neurons responsive to oriented gratings of any orientation yielded the density of orientation-responsive neurons. Multiplying this density by the peak firing rate of each neuron and by the fraction of spikes generated by a single orientation produced the number of spikes elicited in the highest activated voxels in V1 by our

*Correspondence: rafi.malach@weizmann.ac.il
[3]Present address: Center for Neural Science, New York University, New York, NY 10003.

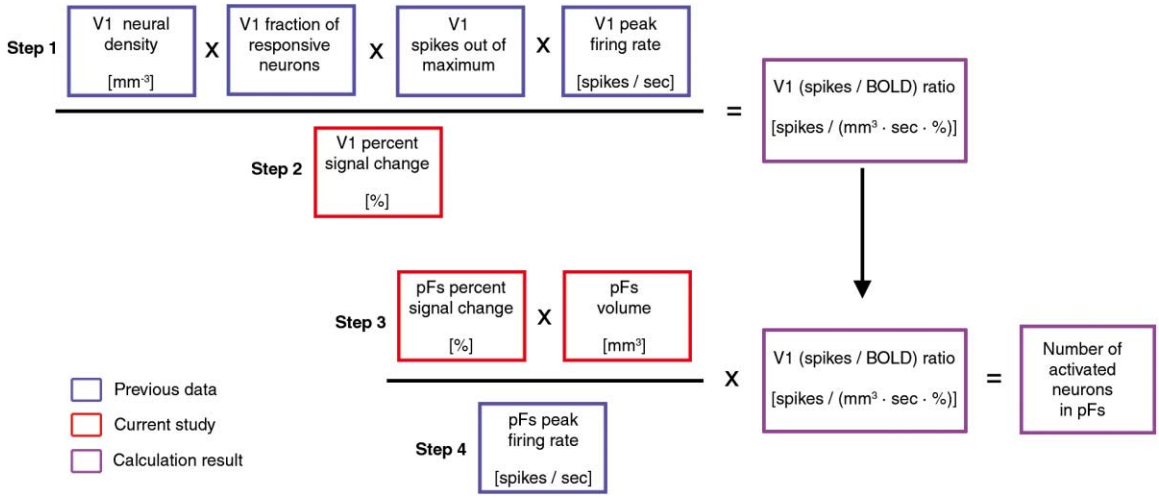

Figure 1. Diagram of the Analysis Steps

Steps one and two are used to calibrate the spikes/BOLD ratio. Steps three and four use this calibration on data from the region of interest (e.g., pFs) to assess a lower bound for the number of responsive neurons in this region.

oriented grating stimulus. This estimate was between 172,800 spikes/s/mm^3 and 810,000 spikes/s/mm^3.

Step Two: V1 BOLD Signal

Voxels in V1, which were most significantly activated by the grating compared to a blank screen, were sampled (Figure 2B). Only highly significant voxels were used to ensure that the receptive field locations of the entire neuronal population in these voxels overlapped the retinotopic location of the stimulus. The average percent signal change across the full hemodynamic response (four time points) was 0.6% ± 0.1% SEM, peaking at 1.0% ± 0.2% (Figure 2B). For comparison, the response in V2 reached a somewhat lower peak (0.7% ± 0.2%).

Combining steps one and two (Figure 1, top) we can conclude that 1% BOLD signal change in V1 roughly reflected between 300,000 and 1,300,000 spikes/s/mm^3.

Step Three: BOLD Response to a Single-Object Image

An external localizer was used to localize the entire extent of visual cortex, as well as face and house-related regions, and the response of each region of interest (ROI) to the single face and house images was measured (see Supplemental Experimental Procedures). In the right visual cortex, a volume of 21,000 ± 5,000 mm^3, located in occipital and temporal regions, showed an average signal change of 0.10% ± 0.04% and 0.14% ± 0.03% for a single face and house, respectively (Figure 2C). In the pFs face-related region, a volume of 700 ± 200 mm^3 was activated at 0.14% ± 0.02% for the face and 0.02% ± 0.03% for the house (Figure 2D, left). Finally, in the CoS building-related region a volume of 500 ± 300 mm^3 exhibited a signal change of 0.26% ± 0.05% for the house and 0.10% ± 0.07% for the face (Figure 2D, right). Left hemisphere activation was sampled in a similar fashion (not shown).

Step Four: Occipito-Temporal Firing Rate

The typical maximal firing rate in high-order object areas was estimated based on monkey inferior temporal (IT), which is considered to be the homolog of human object-related areas [18]. In this area, the typical response is about 20 spikes/s (C.R. Olson, personal communication). However, since representation principles in high-order visual areas are still unclear, it could be that higher firing rates, up to those seen in V1 (45 spikes/s), would have been obtained had the optimal stimuli been used.

Combining steps three and four with the ratio between spike activity and BOLD signal change (Figure 1, bottom) we can now derive an approximation of the number of neurons which responded to single face and house images in the visual cortex in general (Figure 3A) and in the pFs and CoS in particular (Figure 3B). The calculation was done separately in each subject and in each hemisphere. Results for right and left hemispheres were averaged separately across subjects and then summed. Each estimate was calculated twice by using either the most conservative or the most liberal values for each estimate (arrows in Figure 3, see Table 1 for a summary of the estimates).

The results show that a single face image activated between 30 and 300 million neurons in the entire visual cortex, of which between 1.1 and 12 millions were in the pFs face-related region (Figure 3, dark purple bars). A single house image activated between 40 and 400 million neurons, of which between 2.2 and 23 millions were in the CoS (Figure 3, dark blue bars).

To ensure that the results did not depend on the specific manner in which the percent signal change was calculated, we repeated the calculation by using the averaged percent signal change of the two middle time points in each event (Figure 3D, middle bars) or the response peak (Figure 3D, rightmost bars). Very similar results were obtained. Thus we can conclude that at the minimum, a million neurons are activated in object-selective manner even when a single-object image is perceived.

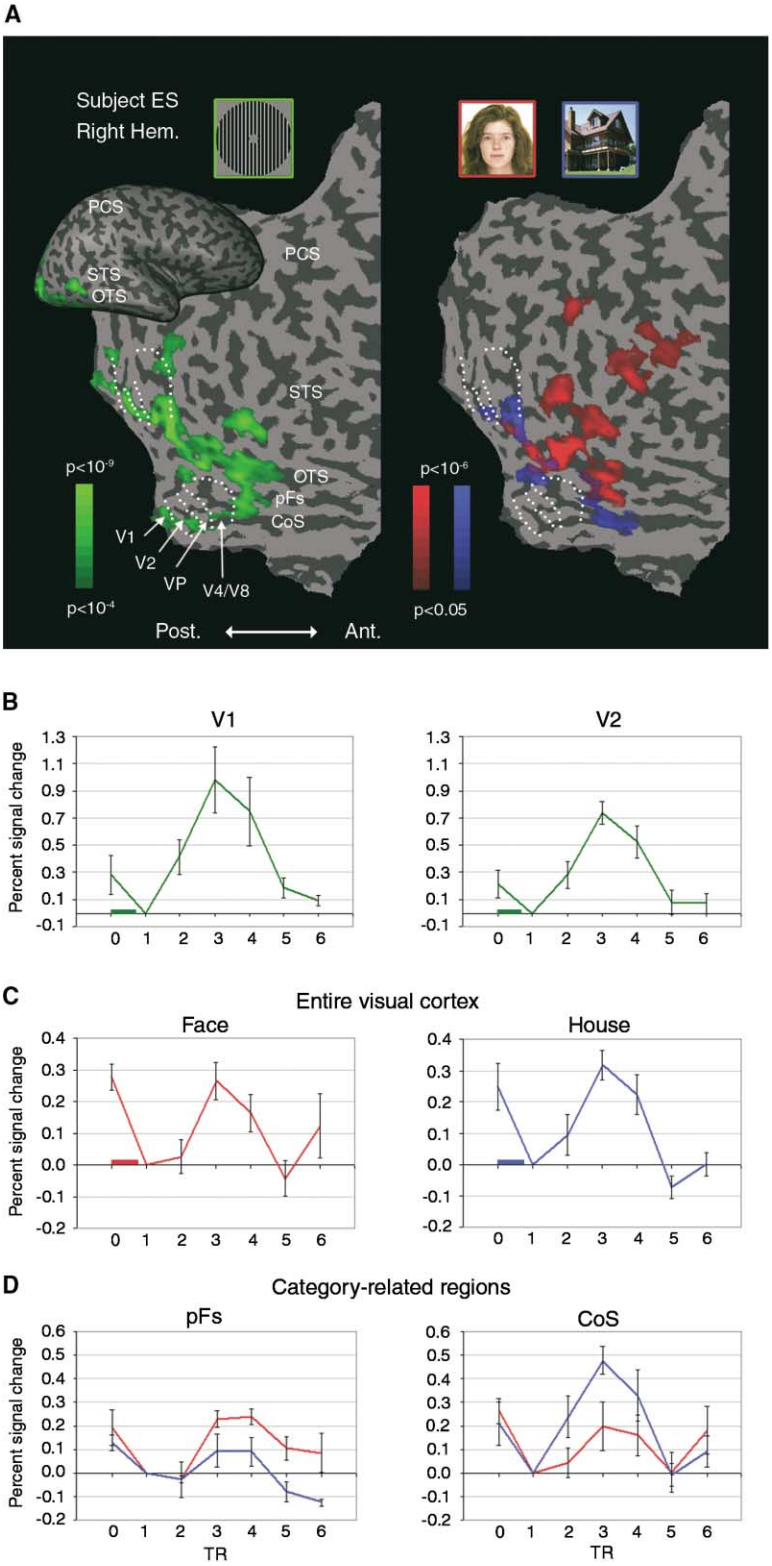

Figure 2. Single-Image Activation Maps and Time Courses

(A) Activation to the moving grating (left) and the face (red) and house (blue) images (right) compared to blank. The right inflated hemisphere is presented from a lateral view for orientation purposes. Dotted lines mark borders of retinotopic areas. Color scales denote statistical significance. Abbreviations: CoS, collateral sulcus; OTS, occipito-temporal sulcus; PCS, post central sulcus; pFs, posterior fusiform gyrus; STS, superior temporal Sulcus; Ant, anterior; and Post, posterior. Note remarkably wide distribution of activation to single-face images and single-house images.

(B) Percent signal change elicited by the grating in V1 and V2.

(C and D) Activation to the face and house images in the entire visual cortex (C) and in the pFs and CoS (D) in the right hemisphere sampled by using an external localizer. Stimuli were presented at TR 0.

Discussion

Our calculation revealed that at least a million neurons in object-related areas and at least 30 million neurons in the entire visual cortex might be activated in response to a single image.

How reliable is the number we obtained? The estimation hinges on several intervening quantities, each exhibiting a rather wide margin of error. Although such errors are minor when compared to the huge number of neurons that turn out in the calculations, we always preferred to err on the conservative side. Thus, a similar

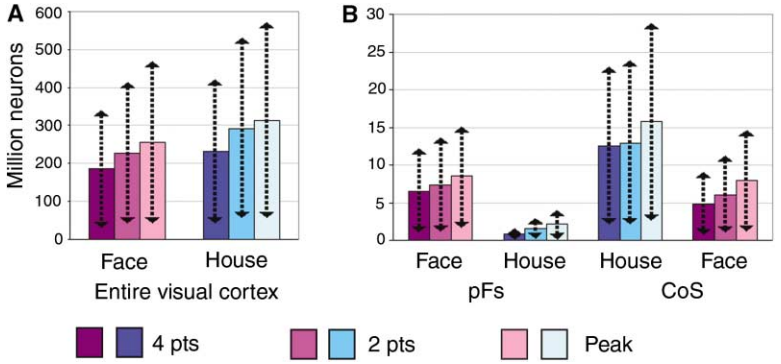

Figure 3. Estimates of the Number of Neurons Responding to a Single Image

(A) Entire visual cortex. (B) pFs and CoS. Estimates are based on percent signal change calculated in three ways (see Supplemental Experimental Procedures). Arrows denote lower and higher estimates using either the most conservative or the most liberal parameter values. Note that a single-face image activates more than a million neurons in human face-selective cortex.

hemodynamic response in V1 and in other visual areas was assumed, although blood vessels are denser in V1 [19], which means that a similar neuronal activation can give rise to a higher hemodynamic response in V1 than in occipito-temporal cortex.

Another implicit assumption that should be noted is a similar excitation-to-inhibition ratio for the grating and the image stimuli and for V1 and object-related areas. However, since more than 80% of cortical neurons are excitatory [20], we can assume that most of the BOLD signal is driven by excitatory rather than inhibitory activity.

The model uses an estimation of the peak firing rate of the neurons in the studied area. This rate may change due to task demands and attentional effects [21–23]. An increase in firing rate will give rise to an increase in BOLD signal and therefore to an erroneous increase in the estimation of responsive neurons. However, this change in firing rate is around 20% on average [21, 22], substantially less than the change we allowed in our estimate of the firing rate (20–45 spikes/s—an increase of more than 100%). Such an increase in firing rate is well within the error limits of the estimation. Furthermore, attentional effects in area V1 will offset this erroneous overestimation, so the error due to attention will actually be caused by a differential attention effect that may exist between V1 and object-related cortex.

Finally, the results are compatible with a recent study showing that the brain's normal energy consumption allows a ceiling of about 30 million concomitantly active neurons [24]. In the future, as more accurate approximations are obtained for each of the above steps, a more and more refined estimate of the number of neurons will be possible.

Table 1. Values of Previous Data Used for the Estimates

Parameter	Value(s)
V1 neural density (neurons/mm^3)	60,000
V1 fraction of responsive neurons	0.2–0.5
V1 spikes out of maximum	0.32–0.60
V1 peak firing rate (spikes/s)	45
Occipito-temporal peak firing rate (spikes/s)	20–45

Each estimate was calculated twice by using either the more conservative values (leading to a lower number of activated neurons) or the more liberal values.

Limitations of the BOLD Signal

So far we have assumed a linear relationship between the BOLD signal and the spike firing rate. Such a relationship has been documented in several studies [4, 5, 25, 26]. However, by using simultaneous recordings of neural signals and fMRI responses, Logothetis et al. [6] have shown that the BOLD signal is explained better by changes in local field potentials (LFPs) than by changes in spike activity. Their interpretation of the result was that the BOLD signal reflects the synaptic inputs to a given area rather than its spiking output. Under normal conditions the synaptic inputs and the spike output are usually highly correlated so that this distinction has no significant practical implications. In fact, Logothetis et al. [6] have shown that for short stimuli, as in our case, an almost linear relationship between spike rate and BOLD signal does exist. The spikes-to-BOLD ratio in V1 calculated here is higher than that reported previously [4], but since this ratio is based on the typical BOLD levels obtained in each lab, the crucial point is that activation in both V1 and occipito-temporal cortex was estimated in the same scan.

However, the fact that the BOLD signal is explained better by LFPs than by spiking activity may have consequences in cases where there is dissociation between the synaptic input and the spiking output, such as when the input is inhibitory [27]. In such cases, the BOLD signal will reflect spiking activity in other brain regions, providing bottom-up and top-down input to the studied area rather than spiking activity in the area itself.

It could be argued that the large number of object-selective neurons estimated by this study is due to the low temporal resolution of the fMRI. Thus, it is conceivable that there may be a rapid process in which the neuronal representation starts out as a representation of local object features, whereas at a later stage, only a few islands of more "holistic" object-selective neurons remain. The sluggish response of the fMRI averages temporal events over seconds and, therefore, may lead to an exaggerated object-representation estimate based on such early, feature-based representation. However, a recent study in our lab [28] has addressed this point directly. By using backward masking, it was demonstrated that holistic, object-completion effects are present right at the very early stages of building up the neural representations, indicating that the neuronal population is object selective right from the initial stages of the response. Similar results were reported by Grill-Spector et al. [29]. Still, the integration of several seconds of

neural activity by the hemodynamic response must be kept in mind when applying a similar approach to assessing the number of participating neurons in other cognitive states.

Theoretical Implications

Our results bear important implications for "realistic" models of object representation. We found that out of about 50 million neurons in the pFs, at least a million (2%) respond to a single face image. Assuming an independent response of each neuron, this, in principle, enables the storage of more than $10^{2,000,000}$ different images (the number of ways to pick one million out of 50 millions). On one hand this approximation may seem exaggerated, since in reality firing of adjacent neurons is slightly correlated [30]. However, on the other hand the approximation is rather conservative, since it assumes an all-or-none neuronal response and a fixed size of the activated subset. Thus, from a combinatorial point of view this representation maximizes the memory capacity of the network leading to a truly stupendous number of possible representations. In contrast, sparse coding (using a representation of very few neurons) or a completely distributed one (using most of the neuronal population) would result in a largely reduced storage capacity.

Could the large number of activated neurons we find serve as a mechanism for enhancing signal to noise? This could be done by way of pooling a large population of neurons having similar receptive field properties and averaging their responses, thus averaging out random "noisy" fluctuations. However, data from single-neuron recordings in monkey cortex argue that such pooling does not help. The reason for that is that the random noise fluctuations appear to be correlated across neighboring neurons so that pooling enhances both signal and noise [31].

Finally, it could be argued that within the highly distributed neuronal population, only the few neurons that are at the very peak of activation truly participate in each representation. At this point, we cannot rule out this possibility. It should be noted though that the population of highly active neurons might be even more widespread than that defined by conventional fMRI, since recent adaptation studies suggest that even weakly activated voxels may contain highly active populations of neurons [32].

Conclusion

A careful estimate shows that at least a million neurons in object-related areas and at least 30 million neurons in the entire visual cortex are activated by a single-object image. This estimate has important implications for theoretical models of object representation, since it implies an enormous cortical storage capacity.

Supplemental Data

Supplemental Data including Experimental Procedures are available at http://www.current-biology.com/cgi/content/full/14/11/996/DC1/.

Acknowledgments

We thank Amos Arieli, Galia Avidan, Galit Fuhrmann, Sharon Gilaie-Dotan, Roy Mukamel, and Yuval Nir for very fruitful discussions. We thank Michal Harel, Eli Okon, and Ida Sivan for technical assistance. This study was funded by ISF 644/99 grant and the Weizmann Institution Center of Excellence. I.L. is funded by a fellowship from the Horowitz Foundation.

Received: January 18, 2004
Revised: March 24, 2004
Accepted: March 24, 2004
Published: June 8, 2004

References

1. Konorski, J. (1967). Integrative Activity of the Brain: An Interdisciplinary Approach (Chicago: University of Chicago Press).
2. Shadlen, M.N., and Newsome, W.T. (1998). The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. J. Neurosci. *18*, 3870–3896.
3. Boynton, G.A., Engel, S.A., Glover, G., and Heeger, D. (1996). Linear systems analysis of functional magnetic resonance imaging in human V1. J. Neurosci. *16*, 4207–4221.
4. Heeger, D.J., Huk, A.C., Geisler, W.S., and Albrecht, D.G. (2000). Spikes versus BOLD: what does neuroimaging tell us about neuronal activity? Nat. Neurosci. *3*, 631–633.
5. Rees, G., Friston, K., and Koch, C. (2000). A direct quantitative relationship between the functional properties of human and macaque V5. Nat. Neurosci. *3*, 716–723.
6. Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., and Oeltermann, A. (2001). Neurophysiological investigation of the basis of the fMRI signal. Nature *412*, 150–157.
7. Tootell, R.B., Hadjikhani, N.K., Vanduffel, W., Liu, A.K., Mendola, J.D., Sereno, M.I., and Dale, A.M. (1998). Functional analysis of primary visual cortex (V1) in humans. Proc. Natl. Acad. Sci. USA *95*, 811–817.
8. Kanwisher, N., McDermott, J., and Chun, M.M. (1997). The fusiform face area: a module in human extrastriate cortex specialized for face perception. J. Neurosci. *17*, 4302–4311.
9. Epstein, R., and Kanwisher, N. (1998). A cortical representation of the local visual environment. Nature *392*, 598–601.
10. Hasson, U., Harel, M., Levy, I., and Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. Neuron *37*, 1027–1041.
11. Pakkenberg, B., and Gundersen, H.J. (1997). Neocortical neuron number in humans: effect of sex and age. J. Comp. Neurol. *384*, 312–320.
12. Hubel, D.H., and Wiesel, T.N. (1968). Receptive fields and functional architecture of monkey striate cortex. J. Physiol. (Lond.) *195*, 215–243.
13. O'Keefe, L.P., Levitt, J.B., Kiper, D.C., Shapley, R.M., and Movshon, J.A. (1998). Functional organization of owl monkey lateral geniculate nucleus and visual cortex. J. Neurophysiol. *80*, 594–609.
14. Ringach, D.L., Shapley, R.M., and Hawken, M.J. (2002). Orientation selectivity in macaque V1: diversity and laminar dependence. J. Neurosci. *22*, 5639–5651.
15. Malach, R. (1994). Cortical columns as devices for maximizing neuronal diversity. Trends Neurosci. *17*, 101–104.
16. Bartfeld, E., and Grinvald, A. (1992). Relationships between orientation preference pinwheels, cytochrome oxidase blobs and ocular dominance columns in primate stirate cortex. Proc. Natl. Acad. Sci. USA *89*, 11905–11909.
17. Geisler, W.S., and Albrecht, D.G. (1997). Visual cortex neurons in monkeys and cats: detection, discrimination, and identification. Vis. Neurosci. *14*, 897–919.
18. Grill-Spector, K., Kushnir, T., Edelman, S., Avidan, G., Itzchak, Y., and Malach, R. (1999). Differential processing of objects under various viewing conditions in the human lateral occipital complex. Neuron *24*, 187–203.
19. Zheng, D., LaMantia, A.S., and Purves, D. (1991). Specialized

vascularization of the primate visual cortex. J. Neurosci. *11*, 2622–2629.

20. Peters, A., and Jones, E.G. (1984). Classification of cortical neurons. In Cereb. Cortex, Cellular Components of the Cereb. Cortex, A. Peters and E.G. Jones, eds. (New York: Plenum), pp. 107–121.
21. Treue, S., and Martinez Trujillo, J.C. (1999). Feature-based attention influences motion processing gain in macaque visual cortex. Nature *399*, 575–579.
22. Spitzer, H., Desimone, R., and Moran, J. (1988). Increased attention enhances both behavioral and neuronal performance. Science *240*, 338–340.
23. Miller, E.K., Li, L., and Desimone, R. (1991). A neural mechanism for working and recognition memory in inferior temporal cortex. Science *254*, 1377–1379.
24. Lennie, P. (2003). The cost of cortical computation. Curr. Biol. *13*, 493–497.
25. Hyder, F., Rothman, D.L., and Shulman, R.G. (2002). Total neuroenergetics support localized brain activity: implications for the interpretation of fMRI. Proc. Natl. Acad. Sci. USA *99*, 10771–10776.
26. Smith, A.J., Blumenfeld, H., Behar, K.L., Rothman, D.L., Shulman, R.G., and Hyder, F. (2002). Cerebral energetics and spiking frequency: the neurophysiological basis of fMRI. Proc. Natl. Acad. Sci. USA *99*, 10765–10770.
27. Mathiesen, C., Caesar, K., Akgoren, N., and Lauritzen, M. (1998). Modification of activity-dependent increases of cerebral blood flow by excitatory synaptic activity and spikes in rat cerebellar cortex. J. Physiol. *512*, 555–566.
28. Lerner, Y., Harel, M., and Malach, R. (2004). Rapid completion effects in human high-order visual areas. Neuroimage *21*, 516–526.
29. Grill-Spector, K. (2003). The functional organization of the ventral visual pathway and its relationship to object recognition. In Attention and Performance XX: Functional Brain Imaging of Visual Cognition, N. Kanwisher and J. Duncan, eds. (London: Oxford University Press), pp. 169–193.
30. Gawne, T.J., and Richmond, B.J. (1993). How independent are the messages carried by adjacent inferior temporal cortical neurons? J. Neurosci. *13*, 2758–2771.
31. Zohary, E., Shadlen, M.N., and Newsome, W.T. (1994). Correlated neuronal discharge rate and its implications for psychophysical performance. Nature *370*, 140–143.
32. Avidan, G., Hasson, U., Hendler, T., Zohary, U., and Malach, R. (2002). Analysis of the neuronal selectivity underlying low fMRI signals. Curr. Biol. *12*, 964–972.