nature neuroscience

A temporal frequency–dependent functional architecture in human V1 revealed by high-resolution fMRI

Pei Sun[1], Kenichi Ueno[1], R Allen Waggoner[1], Justin L Gardner[1,2], Keiji Tanaka[1] & Kang Cheng[1]

Although cortical neurons with similar functional properties often cluster together in a columnar organization, only ocular dominance columns, the columnar structure representing segregated anatomical input (from one of the two eyes), have been found in human primary visual cortex (V1). It has yet to be shown whether other columnar organizations that arise only from differential responses to stimulus properties also exist in human V1. Using high-resolution functional magnetic resonance imaging, we have found such a functional architecture containing domains that respond preferentially to either low or high temporal frequency.

Neurons with similar selectivity for visual stimulus properties, such as orientation and spatial frequency, are grouped in a columnar organization in cat and monkey V1. Neurophysiological[1,2], psychophysical[3] and modeling[4] studies suggest that preference for temporal frequency, an important determinant of how moving images are perceived, also forms a cortical map. Spatio-temporal frequency maps in cats[1] have been demonstrated to contain spatially segregated domains that are selective for a combination of spatial and temporal frequencies, but identical domains (albeit slightly less clearly) have also been found by manipulating temporal frequency alone. Temporal frequency is one of three separable tuning properties proposed in a model[4] based on a recent optical imaging study predicting population responses to moving stimuli in ferrets[5], with orientation and spatial frequency being the other two properties. However, no organization specific to temporal frequency has been demonstrated in primate V1.

In humans, it is not even known whether columns that are segregated as a result of differences in the processing of visual stimulus properties exist in V1. The only columnar architecture that has been demonstrated are ocular dominance columns, the spatial organization of which can be visualized by cytochrome oxidase staining of postmortem brains from subjects who have lost one eye many years before death[6]. This organization is special, both because it is the largest, forming alternating and elongated slabs across V1, and because it arises from both anatomical segregation (input to one of the eyes) and differential cortical processing. Columnar organizations as a result of differences in functional properties have been inaccessible with anatomical staining techniques in humans. Recently, however, functional magnetic resonance imaging (fMRI) has been used on normal human subjects[7] to noninvasively replicate the pattern of ocular dominance columns found with cytochrome oxidase staining. Here, we use similar fMRI techniques to ask whether human V1 contains a functional architecture for stimulus temporal frequency.

To resolve columnar architectures using blood oxygenation level–dependent (BOLD) fMRI, our experimental design and analysis took into account many factors, including the geometrical alignment of single voxels to individual columns, the signal-to-noise ratio (SNR), the signal stability and the spatial specificity of the BOLD signal itself. A primary concern was to select a voxel size and three-dimensional orientation of the imaging slice that optimized our ability to measure columnar organization. Ocular dominance columns are 0.8–1.5 mm in humans[6,7], twice as large as they are in monkeys. Because different columnar structures within a species are nearly equal in size, we estimated that the columns that we were interested in are similar in width and extend 2–3 mm through the gray matter. Ideally, the voxel size should be made as small as possible, much smaller than the column size. But decreasing the voxel size has substantial drawbacks: poor SNR and longer scan times. As a compromise, we used a nonisotropic voxel size (in-plane resolution, 0.75×0.75 mm^2; slice thickness, 3 mm) where the slice was oriented parallel to the gray matter[7]. Using this method, columns can be resolved if the slice plane is orthogonal to the columns, but voxels lose this alignment in heavily sulcated cortex. Using subjects with a straight calcarine sulcus that offered a large flat and wide portion of V1 for imaging resolved this issue (**Fig. 1a**).

To achieve high SNR at a submillimeter resolution, all experiments were conducted with a high magnetic field (4 T), using a small surface coil and a multiple-shot echo-planar imaging pulse sequence (8 shots; volume repetition time, 3.2 s; echo time, 25 ms; average flip angle, 45°; matrix size, 256×256) for functional images. A satisfactory spatial SNR in a single echo-planar imaging image ($\sim$50:1) was achieved at this spatial resolution ($\sim$1.69 mm^3). We achieved a further statistical improvement by acquiring images for the same conditions repeatedly with stable signal over a period of $\sim$25 min.

Signal stability during such a long imaging scan can easily be corrupted by head motion and physiological noise (for example, any head movement $>$1 mm can result in a loss of columnar resolution). During the experiment, rigid head motion was monitored and restricted using a bite-bar. Residual head motion was corrected using in-plane registration during postprocessing. In addition, cardiac and respiratory signals were recorded and removed from time-series images in postprocessing (see **Supplementary Methods** online). These two corrections could improve temporal SNR (ratio of the mean signal intensity and s.d. over time) by $\sim$10–15% (ref. 7). With these measures, and using subjects well trained at restraining head motion and maintaining a stable physiological state, the s.d. of the in-plane center-of-mass fluctuation of

[1]Laboratory for Cognitive Brain Mapping, RIKEN Brain Science Institute, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan. [2]Center for Neural Science and Department of Psychology, New York University, 6 Washington Place, New York, New York 10003, USA. Correspondence should be addressed to P.S. (psun@brain.riken.jp).

Received 2 July; accepted 22 August; published online 14 October 2007; doi:10.1038/nn1983

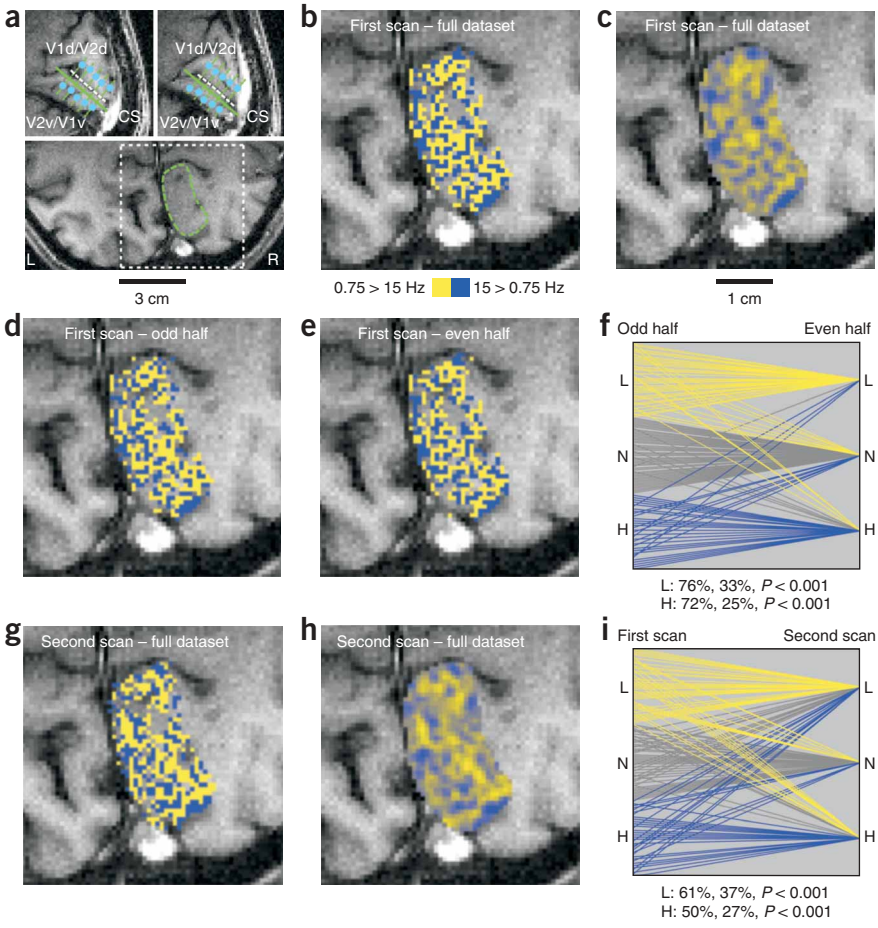

Figure 1 Anatomical details, temporal frequency–domains map and its reproducibility from a representative subject. (**a**) The targeted right V1 is shown in two sagittal images separated by 3 mm (upper panels). The V1/V2 border (blue filled circles) was mapped in five oblique slices (green dashed lines) perpendicular to the calcarine sulcus (CS, white dashed line). The main region of interest (green dashed contour) was in the lower bank of the right CS (lower panel). (**b**) The temporal frequency–domains map generated using the full dataset overlaid on the area indicated by the white dashed box from **a**. (**c**) Smoothed version of **b**. (**d**,**e**) Maps generated by either the odd (**d**) or even (**e**) half of the dataset. (**f**) A parallel-coordinates plot in which the starting point of each line on the left indicates the spatial location of the voxel that it represents (from lower right to upper left in **d**). The color indicates the temporal-frequency preference (yellow, L, low; blue, H, high; gray, N, neither). The ending point on the right indicates the temporal-frequency preference of the corresponding voxel in **e**. For clarity, only the results from every fifth voxel are displayed. Voxels having the same preferences in **d** are grouped together. The percentage of voxels with similar preference across **d** and **e**, the chance percentage and *P* value appear below the plot. (**g**) A map from a second scan in the same session. (**h**) Smoothed version of **g**. (**i**) Parallel-coordinates plot between **b** and **g**.

the time-series images during these experiments was <0.17 mm (**Supplementary Fig. 1** online).

A final major concern that we had about our study was the spatial specificity of the positive BOLD signal that we used. The early, more spatially localized, negative BOLD signal was not used because this signal is both weak and brief, and it has proven to be difficult to use for human studies. The large positive BOLD signal, however, which primarily reflects local cerebral blood flow and volume changes following neuronal activity, has a point-spread function (full width at half maximum, ~2 mm)[8] larger than a cortical column. One effective way to deal with this larger point-spread function is to use complementary stimulus conditions, which elicit maximum cortical responses in one condition and minimum responses in the other. Then, a differential mapping method that compares the two conditions directly can be used to reveal response peaks and troughs at columnar resolution[7,9] (**Supplementary Fig. 2** online).

We used low and high temporal frequencies because neurophysiological findings suggest that they form a pair of complementary stimulation conditions in V1. In cats, nearby V1 neurons prefer similar temporal frequencies[2], and in monkeys, neurons can be divided into groups with low-pass and band-pass temporal tuning characteristics[10,11]. We used a full-screen (~25° × 19°) square wave–modulated white/black (luminance contrast, ~78%) radial checkerboard with a low- (0.75 Hz) or a high-contrast (15 Hz) reversal frequency. Each scan consisted of six 72-volume hyperblocks containing both a 20-volume low-frequency and a 20-volume high-frequency stimulation period, and each period was preceded by 16 volumes acquired with a homogeneous gray screen. To maintain fixation and the level of arousal, the subject was instructed to fixate and to perform a moderately demanding luminance-change detection task on the fixation cross throughout the scan.

We applied a differential mapping method in two steps. First, we generated a pair of activation maps by identifying voxels with a significant response ($P < 0.05$, *t*-test) to low or high temporal frequency compared with a gray screen. The two maps were further thresholded by retaining only voxels whose percent BOLD responses were <5% and by keeping only clusters consisting of at least three connected voxels[7] (**Supplementary Fig. 3** online). These maps defined the spatial extents of activation by the two frequencies in the portion of V1 under study and appeared to be almost identical to each other (**Supplementary Fig. 4** online). Second, we generated a differential temporal frequency–domains map (**Fig. 1b**) by comparing low and high temporal-frequency stimulation directly. Voxels that responded more to low than to high temporal frequency ($P < 0.5$) and responded significantly to low temporal frequency compared with gray screen (as described above) were defined as low temporal frequency–preference voxels, and a similar definition was applied for high temporal frequency–preference voxels. The differential map was sensitized to reveal the peaks and troughs of the signal embedded in the activation maps. As a result, periodically patchy patterns for low (in yellow color) and high (in blue) temporal frequencies can be observed in the differential map. To aid in visualizing the patterns, we smoothed the temporal frequency–domains map (**Fig. 1c**) with a Gaussian filter (full width at half maximum = 0.75 mm).

We evaluated the reproducibility of the obtained temporal frequency–domains map to rule out the possibility that this patchy architecture was simply a result of noise[12]. We grouped the images into independent odd and even halves, and two corresponding temporal frequency–domains maps (**Fig. 1d**,**e**) were generated using the above

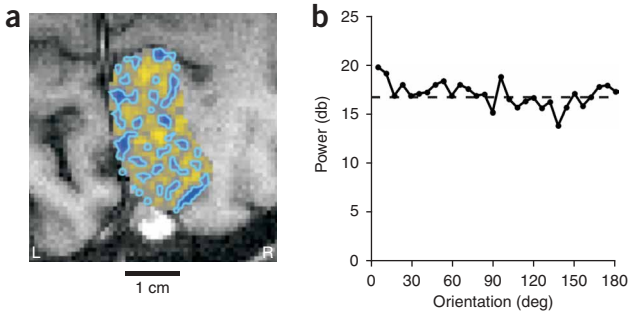

Figure 2 Characteristics of temporal frequency–domains map for the representative subject. (**a**) The edge contours of high temporal-frequency domains (detected using a Marr-Hildreth operator on the smoothed temporal frequency–domains map). (**b**) Power as a function of orientation for the map, showing that the domains were not oriented in any particular direction (Hotelling *t*-square test, compared with equally distributed power spectrum, $P = 0.99$).

procedures. We quantified their similarity by counting how many corresponding voxels in the two maps retained the same preference. We found high percentages of corresponding voxels that had the same preference (76% and 72% of voxels preferred low and high temporal frequency, respectively; **Fig. 1f**), which were significantly above chance (33% and 25%, respectively, both $P < 0.001$, permutation test). Temporal frequency–domains maps from the other three subjects were also highly reproducible (**Supplementary Figs. 5–7** and **Supplementary Table 1** online). In addition, we were able to acquire a second dataset in the same experimental session from two subjects and confirmed intrasession reproducibility (subject 1, **Fig. 1g–i**; subject 2, **Supplementary Fig. 5**).

We noted that the luminance contrast on our display (optic-fiber LCD goggle) was lower for high temporal frequencies because of the slow rise and fall times of image presentation (**Supplementary Fig. 8** online). We conducted two control experiments and found that the peak response evoked by 0.75 Hz was ~94% of that evoked by 15 Hz without contrast adjustment and was ~85% of that evoked by 15 Hz with contrast adjustment (**Supplementary Fig. 8**). To use differential mapping, it is essential to equalize response strength (but not necessarily display contrast; see **Supplementary Fig. 9** online). We therefore opted to use the unadjusted contrast, as it most nearly equalized response strength between high and low temporal frequencies. Our use of unadjusted contrast would not be expected to artificially give rise to a columnar organization, as contrast is generally represented uniformly over the surface of V1. Even if a contrast map does exist, it is substantially weaker than other functional maps[13,14]. Therefore, it is unlikely that the observed patchy patterns in our maps were seriously altered by the contrast difference between the two frequencies.

Compared with ocular dominance columns, the pattern of temporal-frequency domains shows distinct characteristics. Ocular dominance columns form parallel and elongated left- and right-eye stripes, usually orientated perpendicular to the V1/V2 border and interhemispheric fissure[6,7]. In contrast, low temporal-frequency domains appeared to be continuous, whereas the high temporal-frequency domains appeared to be more like isolated islands (**Fig. 2a**), without any particular orientation (Hotelling *t*-square test, $P = 0.99$, compared with equally distributed power spectrum; **Fig. 2b**). We observed similar features from the other three subjects (**Supplementary Fig. 10** online; see also **Supplementary Fig. 9** for discussion). Overall, these features are quite similar to those of cat spatio-temporal frequency domains (low spatial frequency and high temporal frequency versus high spatial frequency and low temporal frequency), which are closely related to the system of cytochrome oxidase blobs and interblobs in V1 (ref. 1). With living human subjects, obviously, we are not able to directly assess the relationship between mapped temporal-frequency domains and cytochrome oxidase blobs and interblobs.

Recently, a single spatio-temporal energy model was applied to explain cortical responses to moving images[5]. However, separable cortical maps for orientation, spatial frequency and temporal frequency may exist and serve as initial filters in the model[4]. Our results demonstrate separate domains in human V1 responding preferentially to low and high temporal frequencies. The existence of temporal-frequency domains in humans also fills a gap in our understanding of physiological processes and human perception of moving images. In cats, two distinct sets of domains processing spatial and temporal frequency differentially have been reported[1]. In primates, the retino-geniculo-striate system is mainly composed of two parallel streams, the parvocellular and magnocellular, which convey and process visual information of different temporal and spatial frequencies[15]. In humans, the existence of two temporal-frequency channels has been suggested in psychophysical studies[3]. The results in this study provide direct physiological evidence that different temporal frequencies are preferentially processed by spatially segregated streams in human V1.

Note: Supplementary information is available on the Nature Neuroscience website.

ACKNOWLEDGMENTS

We are indebted to our subjects for their time and efforts participating in the experiment. P.S. and J.L.G. were supported in part by postdoctoral fellowships from the Japan Society for the Promotion of Science.

Published online at http://www.nature.com/natureneuroscience
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions

1. Shoham, D., Hubener, M., Schulze, S., Grinvald, A. & Bonhoeffer, T. *Nature* **385**, 529–533 (1997).
2. DeAngelis, G.C., Ghose, G.M., Ohzawa, I. & Freeman, R.D.J. *J. Neurosci.* **19**, 4046–4064 (1999).
3. Hammett, S.T. & Smith, A.T. *Vision Res.* **32**, 285–291 (1992).
4. Baker, T.I. & Issa, N.P. *J. Neurophysiol.* **94**, 775–787 (2005).
5. Basole, A., White, L.E. & Fitzpatrick, D. *Nature* **423**, 986–990 (2003).
6. Horton, J.C., Dagi, L.R., McCrane, E.P. & de Monasterio, F.M. *Arch. Ophthalmol.* **108**, 1025–1031 (1990).
7. Cheng, K., Waggoner, R.A. & Tanaka, K. *Neuron* **32**, 359–374 (2001).
8. Shmuel, A., Yacoub, E., Chaimow, D., Logothetis, N.K. & Ugurbil, K. *Neuroimage* **35**, 539–552 (2007).
9. Grinvald, A., Slovin, H. & Vanzetta, I. *Nat. Neurosci.* **3**, 105–107 (2000).
10. Hawken, M.J., Shapley, R.M. & Grosof, D.H. *Vis. Neurosci.* **13**, 477–492 (1996).
11. Foster, K.H., Gaska, J.P., Nagler, M. & Pollen, D.A.J. *J. Physiol.* **365**, 331–363 (1985).
12. Rojer, A.S. & Schwartz, E.L. *Biol. Cybern.* **62**, 381–391 (1990).
13. Carandini, M. & Sengpiel, F. *J. Vis.* **4**, 130–143 (2004).
14. Tootell, R.B., Hamilton, S.L. & Switkes, E. *J. Neurosci.* **8**, 1594–1609 (1988).
15. Van Essen, D.C. & Gallant, J.L. *Neuron* **13**, 1–10 (1994).