

Maps of Visual Space in Human Occipital Cortex Are Retinotopic, Not Spatiotopic

Justin L. Gardner, Elisha P. Merriam, J. Anthony Movshon, and David J. Heeger

Department of Psychology and Center for Neural Science, New York University, New York, New York 10003

We experience the visual world as phenomenally invariant to eye position, but almost all cortical maps of visual space in monkeys use a retinotopic reference frame, that is, the cortical representation of a point in the visual world is different across eye positions. It was recently reported that human cortical area MT (unlike monkey MT) represents stimuli in a reference frame linked to the position of stimuli in space, a “spatiotopic” reference frame. We used visuotopic mapping with blood oxygen level-dependent functional magnetic resonance imaging signals to define 12 human visual cortical areas, and then determined whether the reference frame in each area was spatiotopic or retinotopic. We found that all 12 areas, including MT, represented stimuli in a retinotopic reference frame. Although there were patches of cortex in and around these visual areas that were ostensibly spatiotopic, none of these patches exhibited reliable stimulus-evoked responses. We conclude that the early, visuotopically organized visual cortical areas in the human brain (like their counterparts in the monkey brain) represent stimuli in a retinotopic reference frame.

Key words: reference frame; comparative anatomy; retinotopy; cortex; extrastriate; fMRI; topography; vision

Introduction

Neural responses in early visual areas are believed to be encoded in a retinotopic reference frame: cells respond to stimuli at particular locations on the retina. Although this is a natural reference frame for encoding the location of visual events, other frames of reference may be more suitable for encoding visual stimuli when planning movements (Soechting and Flanders, 1992) or when integrating information from other senses, such as touch or hearing, that are not encoded in the retina (Jay and Sparks, 1984; Stricanne et al., 1996; Groh et al., 2001). Different cortical areas might represent visual stimuli in different frames of reference depending on the function that each area subserves. Determining the reference frame of a visual cortical area would then give an essential clue to its function.

Visual cortical areas are defined by a confluence of factors (physiology, architecture, connections, and topography), but visuotopic mapping with functional magnetic resonance imaging (fMRI) (i.e., retinotopy) has become the standard, routine procedure for identifying human visual areas, noninvasively, in individual subjects (Engel et al., 1994; Sereno et al., 1995; Wandell et al., 2007). It is important to distinguish between measuring a visuotopic map and determining a retinotopic reference frame. Visuotopic maps measure the topographic relationship between

responses in cortex and stimulus positions on the retina with eye position held fixed (typically looking straight forward). Finding an orderly relationship between retinal stimulus position and location in the cortex from such a visuotopic mapping measurement does not necessarily imply a retinotopic reference frame; the reference frame might be based instead on the positions of stimuli relative to the head, body, or other objects in space. Only by systematically changing eye position can one disambiguate alternative reference frames.

Although methods for defining human visual areas with visuotopic mapping have been routine for over a decade, few studies have examined reference frames in human visual cortex. By measuring stimulus-evoked responses for each of several eye positions, d'Avossa et al. (2007) inferred that visual area V1 represents stimuli in a “retinotopic” reference frame, but that visual area MT (also known as V5) (Zeki, 1974; Maunsell and Van Essen, 1983; Tootell et al., 1995a) represents stimuli in a “spatiotopic” reference frame, linked to the location of stimuli in space, independent of eye position.

We used an experimental protocol based on the one described by d'Avossa et al. (2007) to determine the reference frames of 12 human visual cortical areas, each of which was defined with visuotopic mapping. We found that all visual areas, including MT, represent stimuli in a retinotopic reference frame. Our results suggest that the entire visuotopically organized region of the occipital lobe in humans uses a retinotopic reference frame.

Materials and Methods

Subjects. Five healthy subjects (four males) between the ages of 28 and 36 participated in this study. All subjects had normal or corrected-to-normal vision and provided previous written informed consent. Experimental procedures were in compliance with the safety guidelines for MRI research and were approved in advance by the University Committee on Activities Involving Human Subjects at New York University.

Received Dec. 12, 2007; revised Feb. 5, 2008; accepted Feb. 28, 2008.

This work was supported by National Institute of Mental Health Grant R01-MH69880 (D.J.H.). J.L.G. was supported by a Career Award in the Biomedical Sciences from the Burroughs Wellcome Fund and a National Research Service Award (NRSA) from the National Eye Institute (F32-EY016260). E.P.M. was supported by an NRSA from the National Eye Institute (F32-EY016646). We thank the Center for Brain Imaging at New York University for technical assistance.

Correspondence should be addressed to Justin L. Gardner, Department of Psychology, New York University, 6 Washington Place, 8th Floor, New York, NY 10003. E-mail: justin@cns.nyu.edu.

DOI:10.1523/JNEUROSCI.5476-07.2008

Copyright © 2008 Society for Neuroscience 0270-6474/08/283988-12\$15.00/0

Reference frame experiment. We used stimuli similar to the ones used by d'Avossa et al. (2007). In each scan, subjects fixated a cross located at one of three possible screen locations (-10° , 0° , or $+10^\circ$ relative to center) (see Fig. 1A–C, respectively). Visual stimuli appeared in a pseudorandomized order for 3 s at one of four screen locations (-15° , -5° , $+5^\circ$, and $+15^\circ$ from the center of the screen). After each stimulus presentation, the screen was uniform gray (except for the fixation cross) for 3–9 s. Visual stimuli were $1 \times 8^\circ$ patches of 50 vertically moving black and white dots against a uniform gray background. Each dot was 15 arc-min in diameter and moved either up or down at a rate of $15^\circ/\text{s}$. Because these stimuli consisted of high contrast moving dots compared with a uniform gray background, they were well suited for measuring visual responses not only in motion selective areas like MT and V3A but all occipital visual areas we tested. To ensure that subjects maintained fixation at the correct fixation position, we had subjects perform a task at fixation that was continuously adjusted to maintain constant task difficulty (see below), and we monitored eye position for accurate fixation with an infrared eye tracker (see below). For each subject, we acquired either six 8 min (four subjects) or 12 4 min scans (one subject). Subject S2 was scanned twice (on separate days) for the reference frame experiment.

Main experiment data analysis. Data for each functional scan were preprocessed using standard procedures for motion compensation (Nestares and Heeger, 2000), linearly detrended and high-pass filtered with a cutoff frequency of 0.01 Hz to remove low frequency drift, and converted to percentage signal change by dividing the time series of each voxel by its mean image intensity.

Mean hemodynamic response time courses for each stimulus condition were computed using deconvolution (Dale, 1999). That is, the mean response for 25 s after stimulus occurrence was computed by multiplying the pseudoinverse of the stimulus convolution matrix with the time series. This procedure assumes linear temporal summation of the fMRI responses (Boynton et al., 1996; Dale, 1999) but does not assume any particular time course of hemodynamic impulse response functions. To compute the amplitude of response given a canonical hemodynamic response function, we performed a complementary analysis using a general linear model as described below.

The goodness of fit of the deconvolution model, r^2 , was computed as the amount of variance accounted for by the estimated hemodynamic responses (Gardner et al., 2005). That is, the estimated hemodynamic responses computed by deconvolution were convolved with the stimulus times to form a model response time course, and r^2 was then computed as the amount of variance in the original time course accounted for by this model response time course. Statistics (p values) for the r^2 values were computed using a permutation analysis. Event times were randomized and r^2 values were recalculated. We then took 10 of these randomized r^2 distributions computed for all voxels and combined them into a single distribution of r^2 values. We took this combined r^2 distribution as an estimate of the distribution of r^2 values expected by chance. Note that each of the 10 distributions was computed using data from all voxels; thus, combining only 10 randomizations provided sufficient data to estimate the distribution of r^2 expected by chance. Each voxel's p value for each voxel were computed as the percentage of randomized r^2 values that were greater than the actual r^2 value for the voxel.

We used p values computed in this manner to select only voxels that had reliable visual responses for inclusion in the analyses for each visual area. Because each visual area was defined using topographic mapping of the full central visual field (see below), much of the cortex area that we defined was devoted to parts of the visual field in which the stimuli in the main (reference frame) experiment did not appear. We therefore subselected only those voxels for which there was a statistically significant ($p < 0.01$) visual response in at least one of the three eye fixation conditions. This procedure was chosen to ensure that there was a reliable visual response for reference frame testing, but did not make any assumption about what the reference frame was.

To quantify the degree to which a response could be categorized as retinotopic or spatiotopic, we computed a time course index from the deconvolved response time courses. We took conditions that would be expected to match if the responses were in a spatiotopic or retinotopic coordinate frame and computed their difference (see Fig. 5A, pairs of

conditions in gray boxes). We then computed the variance of these residual time courses for retinotopic and spatiotopic predictions separately and computed the index as their difference divided by their sum. For the spatiotopic model, the responses to all four stimulus locations were predicted to match between different fixations, but for the retinotopic model only three or two conditions were predicted to match. To prevent the index from being biased toward a retinotopic outcome by the fewer matching conditions, we selected the following subset of stimulus conditions for the spatiotopic model: the left three stimulus conditions between fixation to the left and center, the right three stimulus conditions between fixation to the right and center, and the center two conditions between left and right fixation. Using different combinations of stimulus conditions for the spatiotopic and retinotopic predictions (e.g., using all matching conditions, or a small subset of conditions that changed the side of fixation they appeared on) made numerical differences to the time course index but did not qualitatively change whether a response was determined to be retinotopic or spatiotopic. The statistical significance of the time course index (e.g., as indicated by the grayscale of the symbols in Fig. 9) was computed using a permutation analysis. Event times were randomized 100 times and the time course index was recomputed for each voxel. The p value for each voxel was then computed as the percentage of the randomized index values that exceeded the actual index value for the voxel.

Response amplitudes were computed by linear regression using a canonical model of the hemodynamic impulse response. The canonical hemodynamic response was a difference of two gamma functions, defined by the following equation: $y = (1/[(a - 1)!])x^{a-1}e^{-x}$, where the shape parameter, a , was set to 6 and 16, respectively, and the second gamma function was scaled by $1/6$. For each of the four stimulus positions, this canonical response function was then convolved with a time course consisting of a 1 when each stimulus occurred and zeroes elsewhere. These four time courses were then placed into the four columns of a design matrix. Each column was then subjected to the same preprocessing as the actual time course data; the columns were linearly detrended, high-pass filtered, and then converted into percentage signal change. The response amplitudes were then computed by taking the pseudoinverse of this design matrix and multiplying it with the actual time course data.

We computed an "amplitude index" using the method in the study by d'Avossa et al. (2007). We estimated the amplitude of response for each eye and stimulus position using a general linear model (see above). These four amplitudes were then linearly interpolated every $1/100$ th of a degree from -20° to $+20^\circ$. The curve for the left fixation was then shifted to the right and the curve for the right fixation was shifted to the left by the same amount, and the summed squared difference between the shifted right and left fixation curves with the central fixation curve was computed. The shift that resulted in the minimum difference between the curves was taken as the amplitude index. By definition, a shift of 1 brings the curves into alignment for a retinotopic reference frame, and a shift of 0 brings the curves into alignment for a spatiotopic reference frame. We evaluated shifts ranging from -0.5 (past the spatiotopic prediction) to 1.5 (past the retinotopic position).

Visuotopic mapping. Visuotopic mapping of visual field eccentricity and polar angle was achieved with standard methods (Engel et al., 1994, 1997; Sereno et al., 1995; Larsson and Heeger, 2006). Briefly, high contrast radial checkerboard patterns were presented either as 90° rotating wedges or as expanding and contracting rings. Each run in a session consisted of 10.5 cycles of length 24 s of the stimulus rotating or expanding/contracting (168 volumes). The first half-cycle of response was discarded. Each scanning session consisted of six runs of the wedge stimulus (three clockwise and three counterclockwise) and four runs of the ring stimulus (two expanding and two contracting). Time series from all runs were advanced by two frames, the response time series for counterclockwise wedges and contracting rings were time-reversed and averaged with responses to clockwise wedges and expanding rings, respectively. The Fourier transforms of the resulting time series were obtained and the amplitude and phase at the stimulus frequency was examined. Coherence was computed as the ratio between the amplitude at the stimulus frequency and the square root of the sum of squares of the amplitudes at all frequencies. We then displayed maps of coherence, amplitude, and phase

on flattened representations of the cortical surface, which were segmented from three-dimensional volumes [T1-weighted magnetization-prepared rapid gradient echo (MPRAGE) volumes, $1 \times 1 \times 1$ mm] using the public domain software SurfRelax (Larsson, 2001). Finally, visual area boundaries were drawn by hand on the flat maps, following published conventions (Larsson and Heeger, 2006), and the corresponding gray matter coordinates were recorded. Because data from each scanning session were registered to the same three-dimensional volume (see below), we could then use these gray matter coordinates to analyze data from subsequent scanning sessions separately for each visual area.

To examine whether human MT exhibited a spatiotopic reference frame, as was reported by d'Avossa et al. (2007), we took particular care in defining MT according to criteria in the published literature. In particular, two primary functional criteria were used to distinguish MT. First, MT responds more strongly to coherently moving dots relative to static dots (Zeki et al., 1991; Tootell et al., 1995a), setting it apart from neighboring areas LO1 and LO2 (Larsson and Heeger, 2006) (see Fig. 2A). Second, MT is identifiable relative to nearby visual areas by its distinct topographic organization (Huk et al., 2002; Smith et al., 2006): MT is located anterolateral to LO1 and LO2 (Larsson and Heeger, 2006), has a representation of the fovea that is different from neighboring visual areas (see Fig. 2B), and has a complete representation of the contralateral hemifield (see Fig. 2C). We further restricted the area defined by these two criteria to include only voxels that modulated with the polar angle (rotating wedge) stimulus (see Fig. 2B). This was done because MT is distinguished from MST in terms of the sizes of the underlying receptive fields (Desimone and Ungerleider, 1986; Tanaka et al., 1986). MST neurons, unlike those in MT, have very large receptive fields that extend into the ipsilateral visual field and are thus not modulated strongly by the polar angle stimulus (Huk et al., 2002; Smith et al., 2006). Applying these criteria enabled us to reliably identify MT in both hemispheres of all five subjects studied.

We note that there is controversy over the definition of area V4 (Tootell and Hadjikhani, 2001; Brewer et al., 2005; Hansen et al., 2007). We used the definition of V4 used by Wandell et al. (2007). Our conclusions would not have differed qualitatively had we adopted any of the other proposed definitions.

Moving dot localizer stimulus. As one of the criteria in defining MT (see above), we compared responses to static and moving dots (Tootell et al., 1995a). A full screen ($29 \times 22^\circ$) pattern of 0.11° square dots appeared on a black background. In one half-cycle of the stimulus, dots moved in an expanding/contracting optic flow pattern whose center was at the fixation point at the center of the screen. On every volume acquisition (1.5 s), we reversed the direction of optic flow to avoid motion aftereffects. On the second half-cycle of the stimulus, the dots remained fixed in location. Each full cycle was 24 s long and the stimulus was run for 11 cycles (176 volumes). Data were acquired from two runs and averaged. The stimulus simulated motion of the observer in three dimensions toward and away from a three-dimensional cloud of dots. Different dots had appropriately computed speeds and directions that varied according to the location of the dot in the simulated three-dimensional volume. Thus, the two-dimensional retinal velocity of each dot varied according to its location in the three-dimensional volume. Across all of the dots presented on the screen, the median retinal speed was $24^\circ/\text{s}$. The two-dimensional projection presented on the screen was maintained to have an average dot density of 5 dots/ $^\circ$. The coherence, phase, and amplitude at the stimulus frequency was computed in the same way as for the topography measurements (see above).

To determine whether the localizer isolated motion-selective responses in MT as opposed to general sensitivity to dynamic versus static patterns, we also measured responses to coherent versus incoherent motion. The coherent motion stimulus was the same as that described above. For the incoherent motion stimulus, each dot was assigned a random optic flow motion on every frame of the stimulus. The random velocities were drawn from the distribution of velocities in the coherent motion stimulus such that local, instantaneous motion was the same, but without global coherence. These stimuli evoked similar responses in MT as the moving/static stimulus; however, early visual areas including V1 did not respond differentially to the coherent/incoherent motion stimuli.

Behavior. To maintain proper fixation and consistent behavioral state during the reference frame experiments, subjects were instructed to perform a 2IFC (two-interval forced-choice) luminance discrimination task on the fixation cross. On every trial of the fixation task, the cross was initially cyan (500 ms), and then briefly dimmed during each of two target intervals (100 ms). The two target intervals were separated by a 500 ms period in which the cross was again cyan. After a final 500 ms cyan interval, the cross turned yellow to indicate the response interval. The subject was given 1 s to press one of two buttons indicating whether the cross was darker during its first or second dimming. If the subject chose correctly, the cross turned green, and otherwise, red. The task was run asynchronously with the peripheral reference frame stimuli. On each trial, the target luminance decrement was set by a two-down one-up staircase to maintain a performance near 71% correct (Wetherill and Levitt, 1965).

MRI methods. MRI data were acquired on a Siemens (Erlangen, Germany) 3T Allegra head-only scanner using a head coil (NM-011; NOVA Medical, Wakefield, MA) for transmit and a four-channel phased array surface coil (NMSC-021; NOVA Medical) for receive. Functional scans were acquired with gradient recalled echo-planar imaging to measure blood oxygen level-dependent (BOLD) changes in image intensity (Ogawa et al., 1990). Functional imaging was conducted with 27 slices placed perpendicular to the calcarine sulcus (repetition time, 1.5 s; echo time, 30 ms; flip angle, 75° ; $3 \times 3 \times 3$ mm; 64×64 gridsize). The first two images were discarded to allow longitudinal magnetization to reach steady-state. A T1-weighted (MPRAGE) $1.5 \times 1.5 \times 3$ mm anatomical volume was acquired in each scanning session with the same slice prescriptions as the functional images. This anatomical volume was aligned using a robust image registration algorithm (Nestares and Heeger, 2000) to a high resolution (three-dimensional MPRAGE, $1 \times 1 \times 1$ mm) volume that was acquired in a separate session.

Visual stimulus presentation. Visual stimuli were presented with one of two displays: (1) an electromagnetically shielded analog liquid crystal display (LCD) flat panel monitor (NEC 2110; NEC, Tokyo, Japan) with a resolution of 800×600 pixels and a 60 Hz refresh rate (S2 and S3 topography experiment) or (2) an LCD projector (Eiki LC-XG100; Eiki, Rancho Santa Margarita, CA) with a pixel resolution of 1024×768 (S1 and S4–S6 topography and all subjects reference frame experiment). The LCD monitor was located behind the scanner bore and was viewed by subjects through a small mirror, at a distance of 150 cm making for a field of view of $16 \times 12^\circ$. Subjects viewed the image from the LCD projector on a rear projection screen placed inside the bore of the magnet at a distance of 57 cm, yielding a field of view of $29 \times 22^\circ$. For the reference frame experiment, our display device could not quite accommodate 15° of visual angle [the angle used in the study by d'Avossa et al. (2007)], so we scaled all of the stimulus and eye positions by a scale factor of 0.9472. Thus, a location of 15° reported in the text was actually scaled to 14.2° , 10° was 9.5° , and 5° was 4.7° . The mirror was positioned so that the stimuli and fixation cross could be viewed with the eyes vertically centered in the orbits, although the stimuli and fixation cross appeared 3.5° below the center of the display. Both display devices were calibrated using a Photo Research (Chatsworth, CA) PR650 SpectraColorimeter to achieve a linear gamma. Stimuli were generated using Matlab with the PsychToolbox [topography experiment (Brainard, 1997; Pelli, 1997)] or MGL [reference frame experiment and MT localizer (<http://justingardner.net/mgl>)].

Eye position measurements. Eye position during the MRI sessions was monitored using an ASL Model 504 eye tracking system (Applied Science Laboratories, Bedford, MA). At the beginning of each functional scan, an eye calibration was done in which the subject fixated a yellow dot at the center as well as 5° to the left, right, above, and below screen center. Data from the calibration were used to find the best affine transformation (translation, rotation, linear scaling, and linear shear) of the raw eye data to eye position in degrees. For most subjects, the difference between the position of the pupil and the corneal reflection was calibrated and used for analysis, but the corneal reflection trace was of poor quality for one subject so only the pupil center position was used. We note that any deviations from fixation would be expected to degrade measurements of a retinotopic reference frame but not of a spatiotopic reference frame (by definition, the responses do not depend on eye position in a spatiotopic

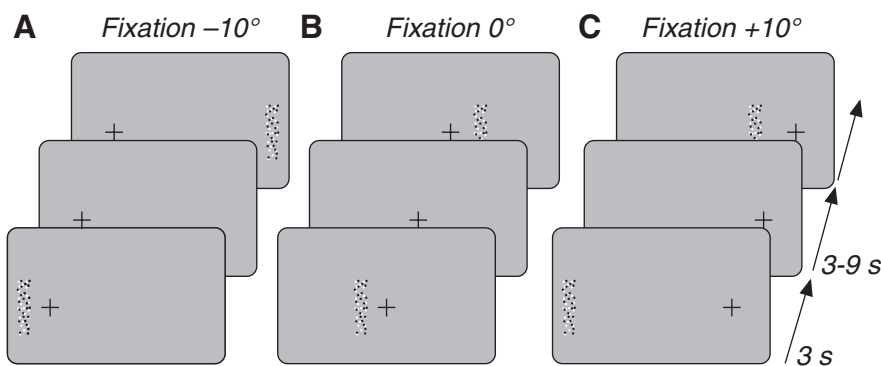


Figure 1. Schematic diagram of experimental conditions used in main reference frame experiment. **A**, **B**, and **C** each depict one of the three different eye fixation conditions that was used during each scan (-10° , 0° , or $+10^\circ$, relative to center, respectively). Visual stimuli, consisting of black and white moving dots, appeared for 3 s, in a pseudorandomized order, at each of four screen locations (-15° , -5° , $+5^\circ$, and $+15^\circ$ from the center of the screen). Between each stimulus presentation the screen was gray for 3–9 s.

representation). Therefore, eye position effects could not explain any lack of spatiotopic response patterns. Nonetheless, we analyzed the eye traces for four of the five subjects for which we had eye tracking data from the main experiment. In particular, we examined the stimulus triggered average of the horizontal and vertical eye position for the 3 s during which the stimulus was presented.

We found that the median eye position during stimulus viewing did not differ as a function of the stimulus position ($p > 0.05$, one-way ANOVA). We also examined the stability of fixation as a function of the three eye positions in the main experiment (for two subjects, we could analyze only two of the three eye positions because the eye tracker failed for one of the eccentric fixation positions). For three of four subjects, the SDs of horizontal and vertical eye positions across trials were $<0.75^\circ$. For the fourth subject, the SDs were $<1.5^\circ$. We suspect that the greater variation in eye position for this subject was attributable to noise in the measurement, because the subject was wearing corrective lenses that interfered with measurement of the corneal reflection, thus forcing us to use only the pupil measurement for eye position. Neither the horizontal ($p = 0.58$) nor the vertical ($p = 0.43$) SD of eye position varied significantly as a function of the three eye positions (one-way ANOVA). For an example of eye position measurements and their analysis, see supplemental Figure 1 (available at www.jneurosci.org as supplemental material).

Results

To test whether human visual cortical areas represent visual stimuli in a retinotopic or spatiotopic reference frame, we measured BOLD fMRI responses to motion stimuli presented at four positions on the screen (-15° , -5° , $+5^\circ$, and $+15^\circ$ from the center of the screen). In separate scans, subjects were instructed to fixate one of three different fixation positions (Fig. 1A–C, -10° , 0° , $+10^\circ$ from the center of the screen, respectively). These stimuli were similar to those used by d'Avossa et al. (2007). To examine responses to these stimuli, we used standard visuotopic mapping procedures to identify visual areas (Engel et al., 1994, 1997; Sereno et al., 1995; Larsson and Heeger, 2006), taking special care to identify area MT according to criteria published by a number of laboratories (Tootell et al., 1995a; Huk et al., 2002; Smith et al., 2006) (see Materials and Methods) (Fig. 2).

V1 uses a retinotopic reference frame. We examined the response time course of the average of all voxels in V1 that responded to the stimuli in at least one of the three eye position conditions ($p \leq 0.01$, permutation test). The left V1 of subject 2 (Fig. 3A) was representative of all subjects, showing responses to stimuli in the visual field contralateral to the fixation point (that is, in retinal coordinates) for all three eye positions.

MT also responded to stimuli in a retinotopic, and not spatiotopic, reference frame. When fixation was held at -10° to the left (Fig. 3B, first column), the three contralateral targets evoked strong responses (red, orange, and blue curves) in MT, and the one ipsilateral stimulus (purple curve) did not. When the subject fixated in the center of the screen (Fig. 3B, second column), the two contralateral targets elicited responses (orange and blue curves). Finally, with the eyes held $+10^\circ$ to the right (Fig. 3B, third column), the single contralateral stimulus elicited a response (blue curve). If MT represented visual stimuli in a spatiotopic reference frame, these stimuli would have evoked similar responses at all eye positions, but plainly the stimuli that evoked maximal responses were at particular locations on the retina. This is the hallmark of a retinotopic

representation. The only readily discernible difference between the data from V1 and MT in Figure 3 is that the selected voxels in MT responded over a wider range of spatial locations, as one might expect from the larger size of MT receptive fields (Desimone and Ungerleider, 1986; Tanaka et al., 1986).

For both V1 and MT, plotting a measure of response amplitude as a function of screen (Fig. 3, fourth column) or retinal (Fig. 3, fifth column) position of the target confirmed the conclusion made from examining time courses; both areas represented stimuli in a retinotopic reference frame. The response amplitudes were calculated using a general linear model, assuming a canonical hemodynamic response function. Response amplitudes across the three fixation conditions were better aligned as a function of position on the retina than as a function of position on the screen.

Like V1 and MT, every visual area that we examined responded in a retinotopic reference frame. Time courses for areas V2, V3, V3A, V3B, V4, V7 (also known as IPS0) (Swisher et al., 2007), LO1, and LO2 showed the same pattern of results as V1 and MT (Fig. 4, plots results from the left hemisphere of subject 2; results were similar for all visual areas in all subjects). Areas with smaller receptive fields behaved like V1, responding to a single contralateral stimulus location that shifted with eye position appropriately for a retinotopic reference frame. Areas with larger receptive fields like V7/IPS0 and LO2 behaved like MT, showing the same characteristic retinotopic pattern of responses to three, two, and one contralateral stimulus locations for the different fixation positions.

To quantify the degree to which responses were described by a retinotopic or spatiotopic reference frame, we computed an index according to the method of d'Avossa et al. (2007) (see Materials and Methods). This index was based on response amplitudes estimated by a general linear model, so we call it the "amplitude index." The amplitude index was computed by finding the best alignment of linearly interpolated curves like the ones in the fourth and fifth columns of Figure 3. Perfect alignment in a retinotopic reference frame would have given a value of 1, and perfect alignment in a spatiotopic reference frame would have given a value of 0.

The amplitude index is only meaningful if the underlying response time courses are well fit by the canonical hemodynamic response function used in the general linear model. To avoid this assumption, we performed a complementary analysis with a second index, the "time course index" that is nonparametric and does not

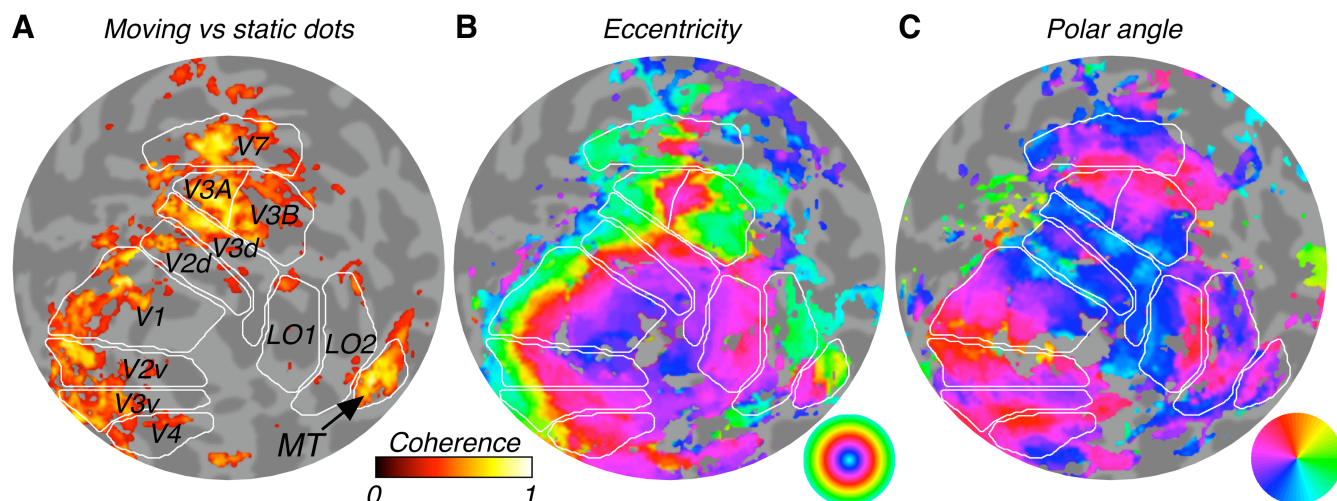


Figure 2. Criteria used to establish the location of visual areas including MT (subject S1, right hemisphere). For each panel, activity is rendered on a flattened representation of the occipital cortex. **A**, Responses to coherently moving dots relative to static dots. The color at each point on the map indicates the coherence of the response at the stimulus frequency. **B**, Map of visual eccentricity. The color at each point represents the response phase of the voxel to an expanding/contracting ring stimulus. **C**, Map of polar angle. The color at each point on the map represents the response phase of the voxel to a rotating wedge stimulus.

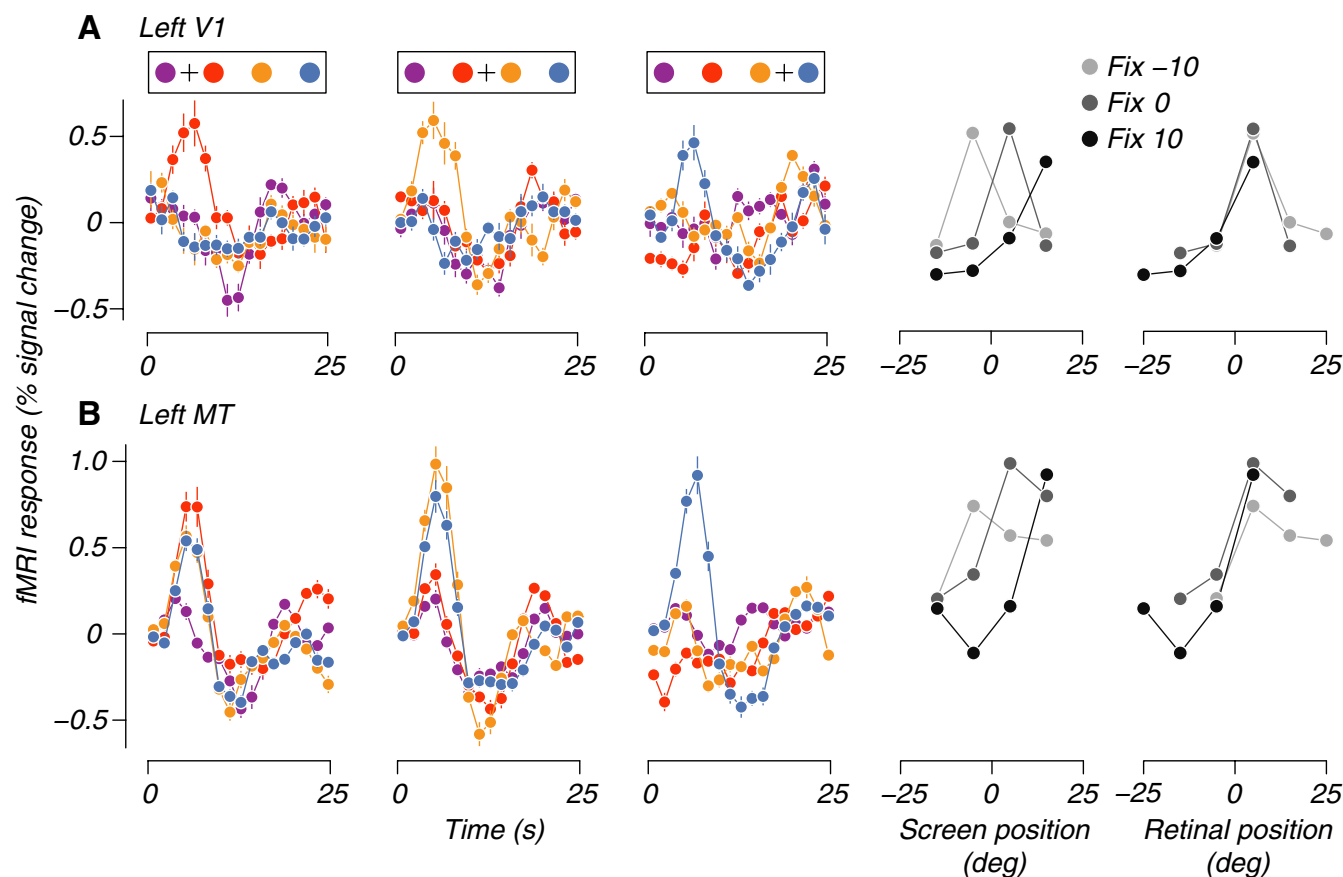


Figure 3. Retinotopic responses from the left hemisphere of a representative example subject (subject S2). **A**, Left hemisphere V1. **B**, Left hemisphere MT. Mean responses were computed for all voxels in which there was a significant visual response in at least one of three fixation conditions ($p < 0.01$, permutation test). Columns 1–3 plot the responses for fixation conditions at -10 , 0 , or $+10^\circ$, relative to the center, respectively. Purple, red, orange, and blue traces show the responses to visual stimuli presented at -15 , -5 , $+5$, and $+15^\circ$ from the center of the screen, respectively (see **A**, inset). Columns 4–5 show the amplitudes of responses as computed by a general linear model using a canonical hemodynamic response, plotted as a function of position on the screen (column 4) or on the retina (column 5). Light gray, dark gray, and black traces represent amplitude of responses for fixation at -10 , 0 , or $+10^\circ$, respectively (see legend).

make any assumptions about the shape of the hemodynamic response. The time course index measures the fraction of the variance in the first 10 s of the BOLD response after stimulus onset that can be accounted for by a retinotopic or a spatiotopic reference frame

model. Specifically, we subtracted pairs of responses from different eye fixations that corresponded to the same retinotopic or spatiotopic position, respectively (Fig. 5A). If the responses were perfectly accounted for by one of the models, the residual variance for

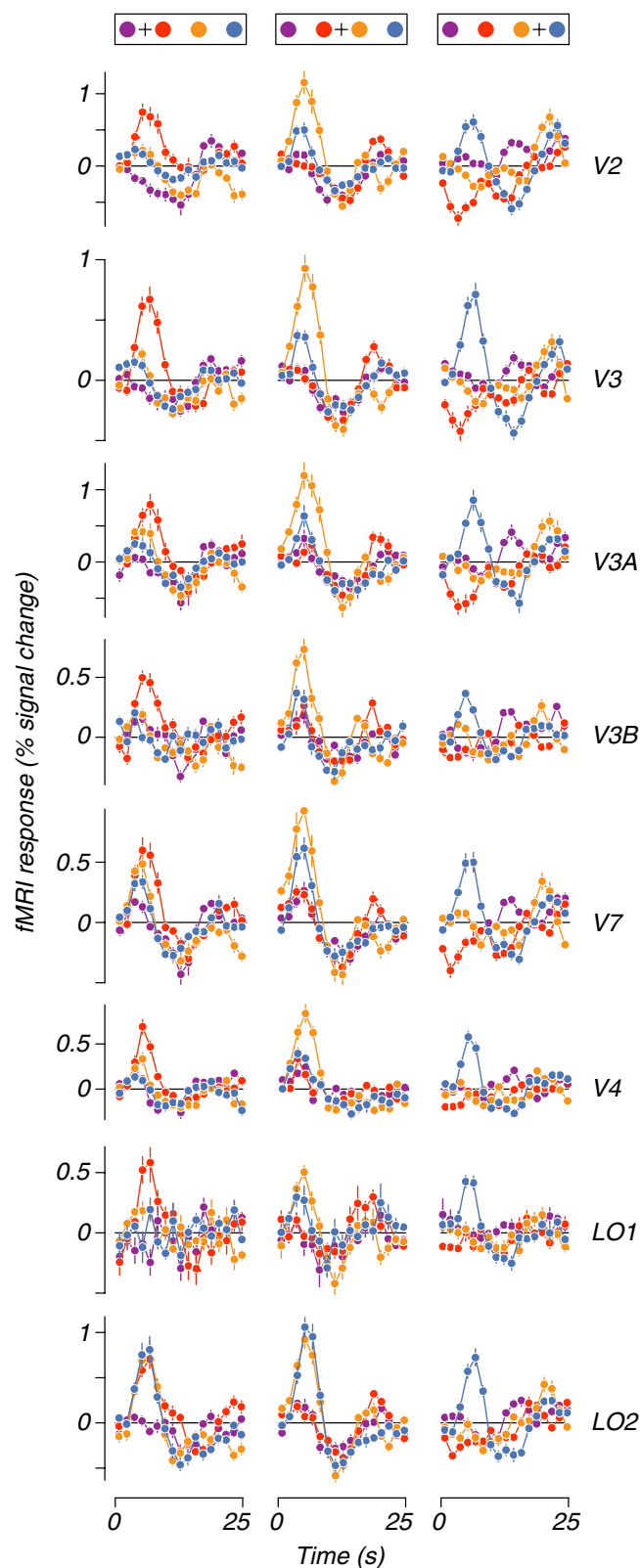


Figure 4. Example of retinotopic responses from left hemisphere V2, V3, V3A, V3B, V4, V7 (also referred to as IPS0), LO1, and LO2 from subject S2. The conventions are the same as in Figure 3.

that model would have been 0. We then computed the index as the difference in residual variances of the spatiotopic and retinotopic models divided by their sum. This index would have been -1 for

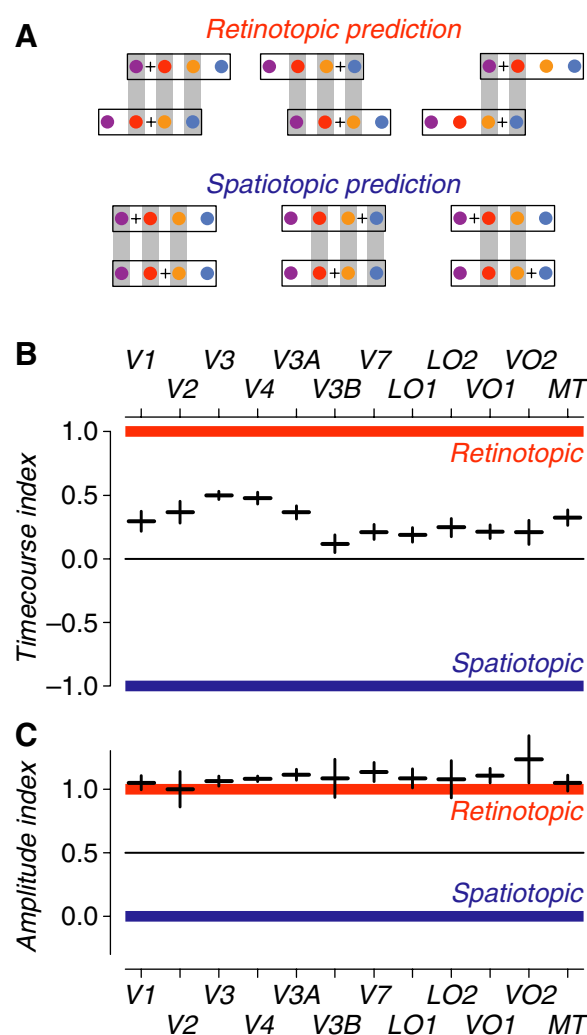


Figure 5. Summary statistics showing retinotopic responses for all visual areas across all subjects. **A**, Tableau depicting the pairs of conditions that would be predicted to display matching responses for a retinotopic reference frame (top row) or a spatiotopic reference frame (bottom row). Although all stimulus pairs should match for the spatiotopic reference frame, the most eccentric stimulus condition was dropped from calculation to have the same number of comparisons as for the retinotopic reference frame (see Materials and Methods). **B**, Time course reference frame index for all visual areas averaged across subjects. Error bars indicate SEM across subjects. The red and blue lines indicate predictions for perfectly retinotopic or spatiotopic responses. **C**, Amplitude reference frame index for all visual areas averaged across subjects.

perfectly spatiotopic and $+1$ for perfectly retinotopic responses. A value of 0 would have indicated that both models explained the data equally well (or equally badly).

Factors unrelated to reference frame, such as noise in the responses, prevented both the amplitude and time course indices from reaching a perfect retinotopic or spatiotopic value. The time course index, which required estimation of all the time points in the response as opposed to a single amplitude measure, was more susceptible to noise than the amplitude index. Furthermore, any systematic differences in the hemodynamic response across eye positions would have caused the indices to deviate from 1. Potential sources of systematic variability included overall changes in response magnitude with eye position, as might be found for neurons with gain fields (Andersen and Mountcastle, 1983; Galletti et al., 1995), and the possibility that at different fixation positions the edge of the projection screen may have provided a stationary luminance edge that contributed to the responses regardless of stimulus location.

Both the time course index (Fig. 5B) and the amplitude index (Fig. 5C) suggest that a retinotopic reference frame describes all visual areas in all subjects (12 hemispheres including one subject who was scanned twice; one or two hemispheres were dropped from the analysis for areas V4, V3B, V7, LO1, and VO1 because of lack of significant visual activity; only two hemispheres in which we were able to visuotopically define and measure significant visual activity from VO2). The amplitude index was statistically indistinguishable from the retinotopic prediction of 1. The time course index also displayed retinotopic values, but did not reach perfect retinotopic values of 1. This was to be expected, however, because the time course index was more susceptible to noise (see above) than the amplitude index. In specific contradiction to the report of d'Avossa et al. (2007), we found that the time course index for area MT was well within the range of retinotopic values displayed by other visual areas.

It is of course possible that even if the average response of a whole area is retinotopic, there might exist subregions within the area that use another reference frame. We therefore examined whether responses on a voxel-by-voxel basis, as opposed to averaged over each visual area, could be classified as being in a retinotopic or spatiotopic reference frame. We plotted either the time course index (Fig. 6B) or the amplitude index (Fig. 6C) on a flattened representation of a posterior region of cortex (depicted in Fig. 6A) that included most of the occipital lobe. We display index values only for voxels that exhibited strong visual responses to at least one of the stimulus conditions in at least one of the three fixation conditions ($p \leq 0.01$, permutation test) (for details, see Materials and Methods). Inspection of the maps of both indexes confirmed that all visual areas examined were primarily retinotopic (red) as opposed to spatiotopic (blue). Although some voxels had responses that could not be classified (white) by these two indexes, none of the voxels on these maps displayed index values that were distinctly spatiotopic.

d'Avossa et al. (2007) reported spatiotopic responses in area MT, which we did not observe. We therefore wondered whether there might be a region of spatiotopic responses near MT in our subjects that might account for the discrepancy between our finding and theirs. We began by making maps of time course indexes (Fig. 7A) and amplitude indexes (Fig. 7B) like the ones shown in Figure 6, but with two modifications. First, to better identify regions near area MT, we shifted the flat patch location so that MT was in the center (patch location depicted in Fig. 7B, bottom, inset). Second, so as not to miss any possible spatiotopic responses, we displayed all voxel indexes, whether they had significant visual activity or not. Visual examination of these maps showed clearly retinotopic indexes (red) inside visual areas (black outlines). Inspection of individual voxels within these clearly retinotopic regions revealed responses like the one shown in Figure 7C. These single voxel time courses had clear visually evoked responses in a retinotopic reference frame, similar to the responses computed for the average over the whole visual area (compare Fig. 3B). Plotting the amplitude of response for these voxels as a function of position on the screen (Fig. 7C, fourth column) or position on the retina (Fig. 7C, fifth column) yielded curves better aligned in retinotopic than spatiotopic coordinates.

Inspection of these index maps revealed ostensibly spatiotopic voxels, but the time courses for these voxels did not exhibit reliable visual activity and were otherwise indistinguishable from noise. Examples of such ostensibly spatiotopic voxels chosen from the amplitude index map are shown in Figure 7, D–F (the voxel locations are indicated in Fig. 7B). We note that there is an obvious selection bias in examining statistics of voxels that have

been chosen to have a spatiotopic amplitude or time course index; by definition, the amplitude or time course must show the spatiotopic effect for which they have been selected. However, spatiotopic values of either index do not necessarily ensure that the voxel time course will have a robust visual response with the usual hemodynamic response. It is therefore useful to examine the time courses of these ostensibly spatiotopic voxels to make sure that they had a credible visual response. Indeed, the time courses for these representative voxels, which included the voxel with the smallest (i.e., most spatiotopic) amplitude index (Fig. 7F), had no clear visually evoked response (first to third columns). Inspecting the estimated response amplitudes on an expanded vertical scale (fourth and fifth columns) showed the reason why these voxels had small index values; the estimated response amplitudes, although mostly because of noise, were better aligned in screen coordinates than retinal coordinates. Similarly, voxels with ostensibly spatiotopic responses chosen from the time course index map (Fig. 7G–H), including the voxel with the smallest (i.e., most spatiotopic) time course index (Fig. 7G), showed no clear visually evoked response. Indeed, all the ostensibly spatiotopic voxels we found in the vicinity of MT resulted from noisy responses that did not reflect visually evoked activity.

We used a cross-validation method to determine whether any spatiotopic responses in and around MT were reliable. For each subject, we split the data in half, and computed the time course index for each half of the data. Based on a histogram of the index values for all voxels in MT+ (i.e., MT plus the neighboring motion sensitive areas including area MST) in the first half of the data, we selected voxels that could be categorized as retinotopic (Fig. 8A, time course index ≥ 0.2 , dark gray tail of distribution). Calculating the time course index on the second half of the data for these retinotopic voxels confirmed that these responses were reliably retinotopic; the distribution of time course indexes for these voxels in the second half of the data were again in the retinotopic range (Fig. 8C). Performing the same analysis for the ostensibly spatiotopic voxels (Fig. 8A, time course index less than or equal to -0.2 , light gray tail of distribution), we found that the responses were not reliable. On the second half of the data, the index values were not in the spatiotopic range, but were centered around 0 (Fig. 8B). In fact, computing the correlation of the time course index for all voxels with a nominally retinotopic index in the first half of the data (time course index > 0) with the index value in the second half of the data, gave a positive correlation ($r = 0.120$; $p < 0.001$), indicating that there was a significant tendency for values to remain retinotopic in the second half of the data. Computing the same correlation for the voxels with a nominally spatiotopic index (time course index < 0), the correlation was significantly negative ($r = -0.215$; $p < 0.001$), indicating that the response in the second half of the data tended not to replicate the spatiotopic effect, but instead were more retinotopic.

Voxels with more reliable visual responses were more clearly classified as being retinotopic. An index value of 0 might have been observed in some voxels for either of two reasons: first, if the responses were completely unreliable (noise only) and, second, if the reference frame was halfway between purely retinotopic and purely spatiotopic. The cross-validation analysis suggests the former (noise) interpretation. To confirm that this was the case, we performed an analysis comparing the time course index with a measure of response reliability. We calculated the reliability of responses, r^2 , by taking the fraction of the variance in the original time course that was accounted for by the mean (technically, deconvolved) response time course (Gardner et al., 2005). A voxel with no repeatable response to the visual stimuli would

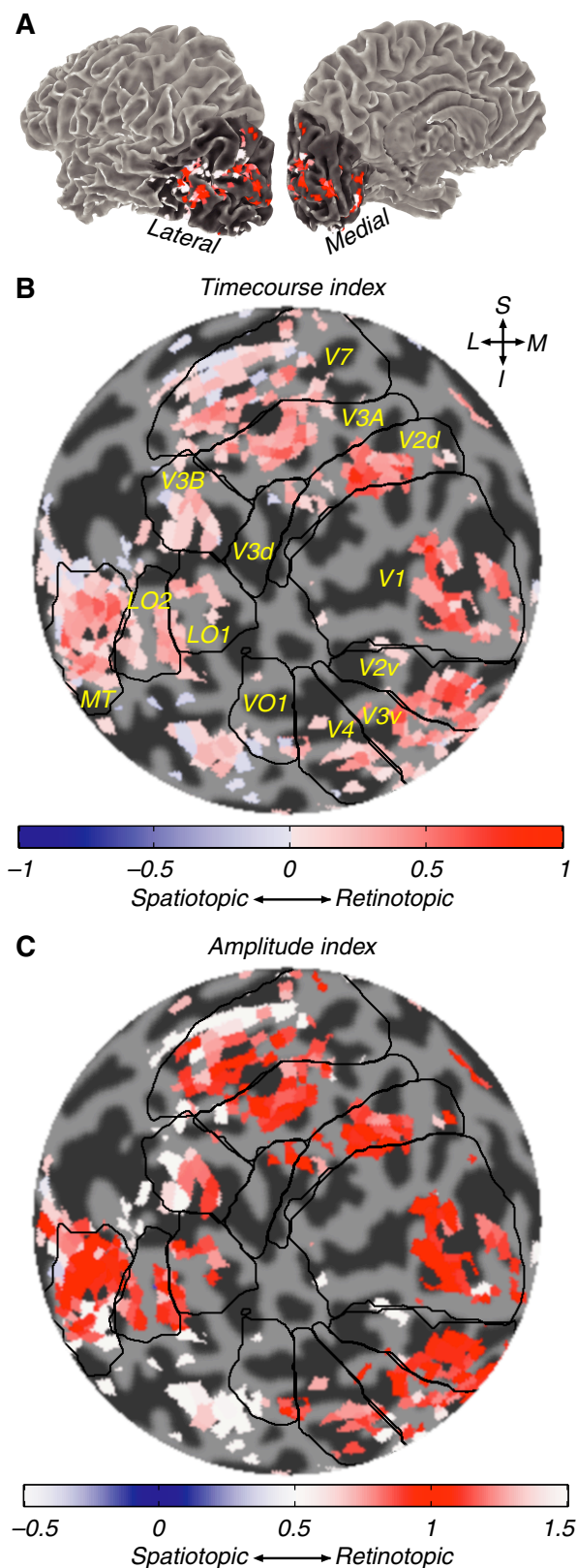


Figure 6. Reference frame maps for the left hemisphere of a representative example subject (subject S2). **A**, View of the lateral (left) and medial (right) surfaces of S2's left hemisphere showing the location of the flattened patches in **B** and **C**. Superimposed on the cortical surface (color) is the amplitude index map. **B**, **C**, Flattened representation of subject S2's occipital cortex with the time course index (**B**) or the amplitude index (**C**) for each voxel that had a significant visual response in at least one of the three fixation conditions ($p < 0.01$, permutation test). The regions in red are retinotopic, and those in blue are spatiotopic. White indicates that reference frame was ambiguous.

have had an r^2 of 0. A voxel whose time course consisted of identical responses to each stimulus presentation would have had an r^2 of 1. Plotting the r^2 value against the time course index for all voxels in MT across all subjects, we found a significant positive correlation between r^2 and the time course index ($p < 0.05$, all hemispheres) (Fig. 9A).

We found this same relationship across all visual areas in all subjects. Robust visual responsiveness was always associated with a retinotopic reference frame. Summarizing plots like the one in Figure 9A, with a mean and SE ellipse for all the voxels in each area, separately for each subject, showed that there was a positive correlation between the robustness of response and the time course index (Fig. 9B) (note change of scale from Fig. 9A). MT (black ellipses), for some subjects, had both highly reliable responses and high (i.e., retinotopic) time course indexes (for subject-by-subject plots, see supplemental Fig. 2, available at www.jneurosci.org as supplemental material). Measurements from other subjects were noisier and consequently had lower r^2 values and less clearly retinotopic time course indexes; this was true of all visual areas, including V1. Although one might expect that the earliest visual areas would be the easiest in which to distinguish the reference frame, we found instead that response reliability was the factor that most strongly determined whether the reference frame was unambiguous. For example, V1 in some subjects had both small r^2 and small index values, whereas V4 in some subjects had both large r^2 and large index values. That is, having a robust visual response was a prerequisite for testing which reference frame that response was in.

Discussion

Although multiple maps of the visual world have been discovered in the human cortex by using visuotopic mapping procedures (Engel et al., 1994; Sereno et al., 1995; Wandell et al., 2007), the reference frame of these visual maps has not been systematically tested. We used these visuotopic mapping procedures to define 12 visual areas that contain maps of the visual world. By measuring activity evoked by different stimulus locations and different eye positions, we have shown that all of these cortical areas, including area MT, represent stimulus location in a retinotopic reference frame.

We searched for spatiotopic responses throughout the occipital lobe, paying extra attention to the region in and around area MT. We found some voxels (i.e., small patches of cortex in individual brains) that appeared to be spatiotopic, but these turned out to be spurious; cross-validation from half of the data to another revealed that only retinotopic responses were reliable. In fact, the reliability of visual response was correlated with the degree to which visual areas could be classified as being retinotopic, and all robust visual responses in our data were clearly more consistent with a retinotopic than a spatiotopic reference frame.

Our results are consistent with experiments in awake behaving monkeys showing that visual cortical areas, and particularly all those with visuotopic maps, represent the world in a primarily retinotopic reference frame (Cohen and Andersen, 2002). The most commonly reported effect of eye position is to change the overall gain of visual responses (Andersen et al., 1985b; Bremmer et al., 1997). Such gain fields scale the magnitude of retinotopic visual responses as a function of the position of the eye in the orbit. The effect can be fit with a planar function of eye position that varies from neuron to neuron. If neurons with similar gain fields are grouped together in the cortex (Siegel et al., 2003), then fMRI responses may similarly show changes in the gain of response for different eye positions (Baker et al., 1999; DeSouza et

al., 2002). In fact, gain field effects could account for part of the reason why our indices did not reach perfect retinotopic values; while responses were in a retinotopic reference frame, they would be scaled differently at different eye positions. To fully test for this type of effect would require considerably more data, ideally measuring the response fields of each voxel for each of several eye positions that vary along the horizontal and vertical dimensions to reliably fit a gain field plane to the responses.

There are some reported exceptions to retinotopic reference frames (Graziano et al., 1994; Olson and Gettner, 1995; Duhamel et al., 1997; Dean and Platt, 2006; McKyton and Zohary, 2007), notably in the ventral intraparietal area (VIP) (Maunsell and van Essen, 1983). VIP is sensitive both to somatosensory stimuli around the head and to visual motion stimuli such as looming optic flow motions, and does not reportedly have a clear topographic organization (Colby et al., 1993; Duhamel et al., 1998). Neurons in VIP display a continuum of responses with approximately equal numbers representing stimuli in retinotopic and spatiotopic reference frames (Duhamel et al., 1997; Avillac et al., 2005). We did not identify human VIP (Serenio and Huang, 2006) in our experiment. Although it has been suggested that putative human VIP uses a spatiotopic reference frame (Serenio and Huang, 2006), the fMRI measurements that we made would have averaged responses from neurons with different reference frames, and thus would not have been likely to reveal the spatiotopic representation of a subpopulation of neurons.

Retinotopic reference frames in early visual areas may be the result of static anatomical connections that link visual responses in cortex to particular parts of the retina. However, there is evidence from single-unit recording studies that retinotopic representations, predominately in higher order areas and ones involved in the generation of saccadic eye movements, may be actively updated in anticipation of the end point of an impending saccade (Goldberg and Bruce, 1990; Duhamel et al., 1992; Walker et al., 1995; Umeno and Goldberg, 1997; Nakamura and Colby, 2002). Human fMRI experiments have found results in agreement with these studies (Medendorp et al., 2003; Merriam et al., 2003, 2007), particularly for regions of the intraparietal cortex, the putative human homolog to monkey LIP (lateral intraparietal area) (Andersen et al., 1985a). Although these and other studies (Tolias et al., 2001) have suggested that visual responses may be actively modified or updated around the time of saccade initiation, they are all consistent with retinotopic

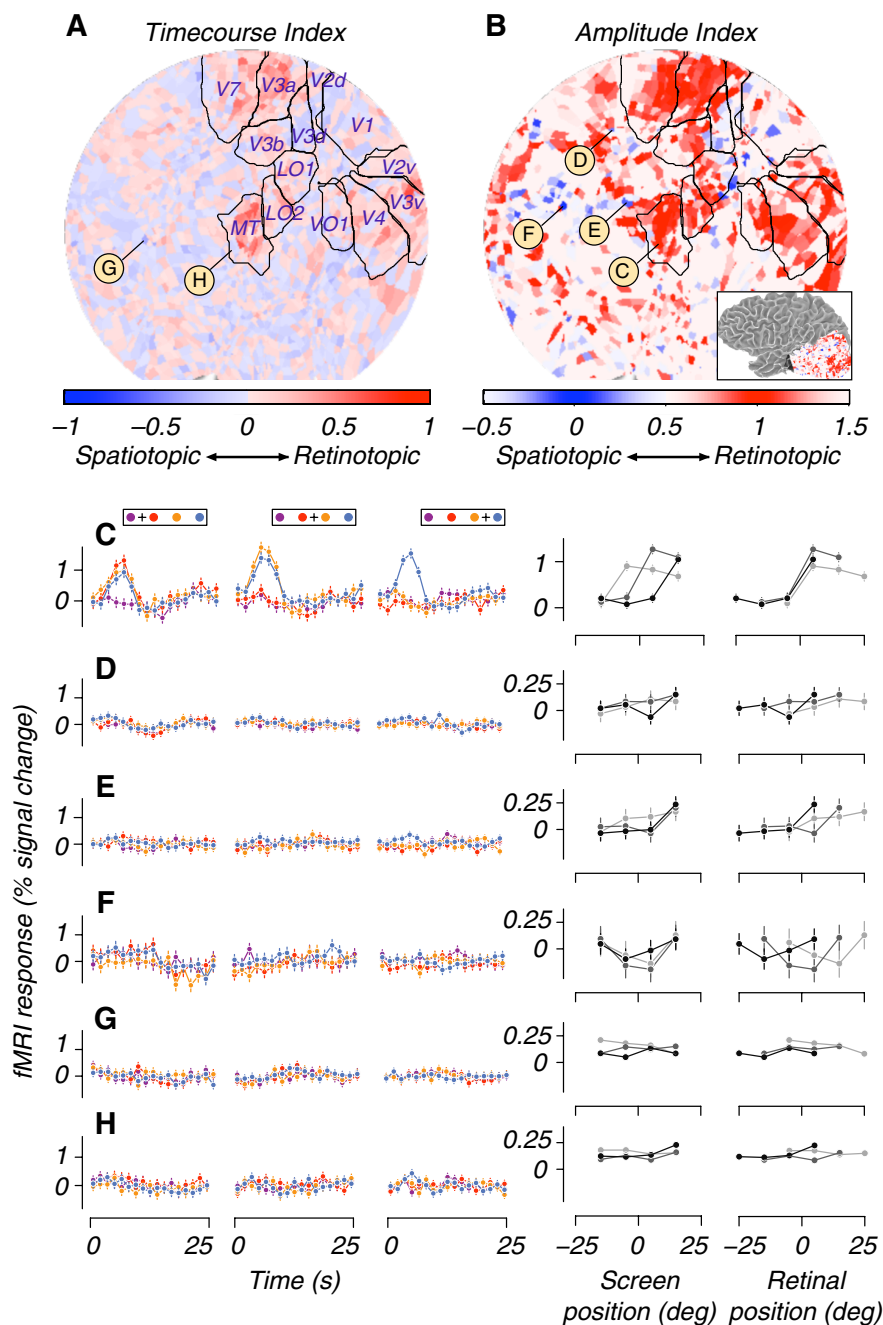


Figure 7. Examination of time courses for ostensibly spatiotopic voxels identified by the time course or amplitude indexes. **A**, Flat map of the time course reference frame index from S2's left hemisphere. The conventions are the same as in Figure 6 except that no statistical threshold was applied to the map. As depicted in the color bar at the bottom of the panel, saturated colors represent voxels with index values close to the ideal spatiotopic or retinotopic value. **B**, Amplitude index. The inset at bottom of **B** shows location of patch on a surface rendering of the left hemisphere with the amplitude index superimposed. **C–H** show responses from individual example voxels (same format as Fig. 3). **C** is an example of a retinotopic response from MT. **D–F** are examples of voxels with ostensibly spatiotopic responses as computed by the amplitude index (**E** has the largest spatiotopic amplitude index found in this hemisphere). **G** and **H** depict examples of ostensibly spatiotopic voxels according to the time course index (**G** has the largest spatiotopic time course index from this hemisphere).

reference frames when the eyes are fixed, as we found in this study.

How can these results be reconciled with the claim that human MT is organized in a spatiotopic reference frame (d'Avossa et al., 2007)? Two factors in our experimental design and analysis were critical. First, we defined MT using visuotopic mapping as well as functional criteria. d'Avossa et al. defined an area based on the

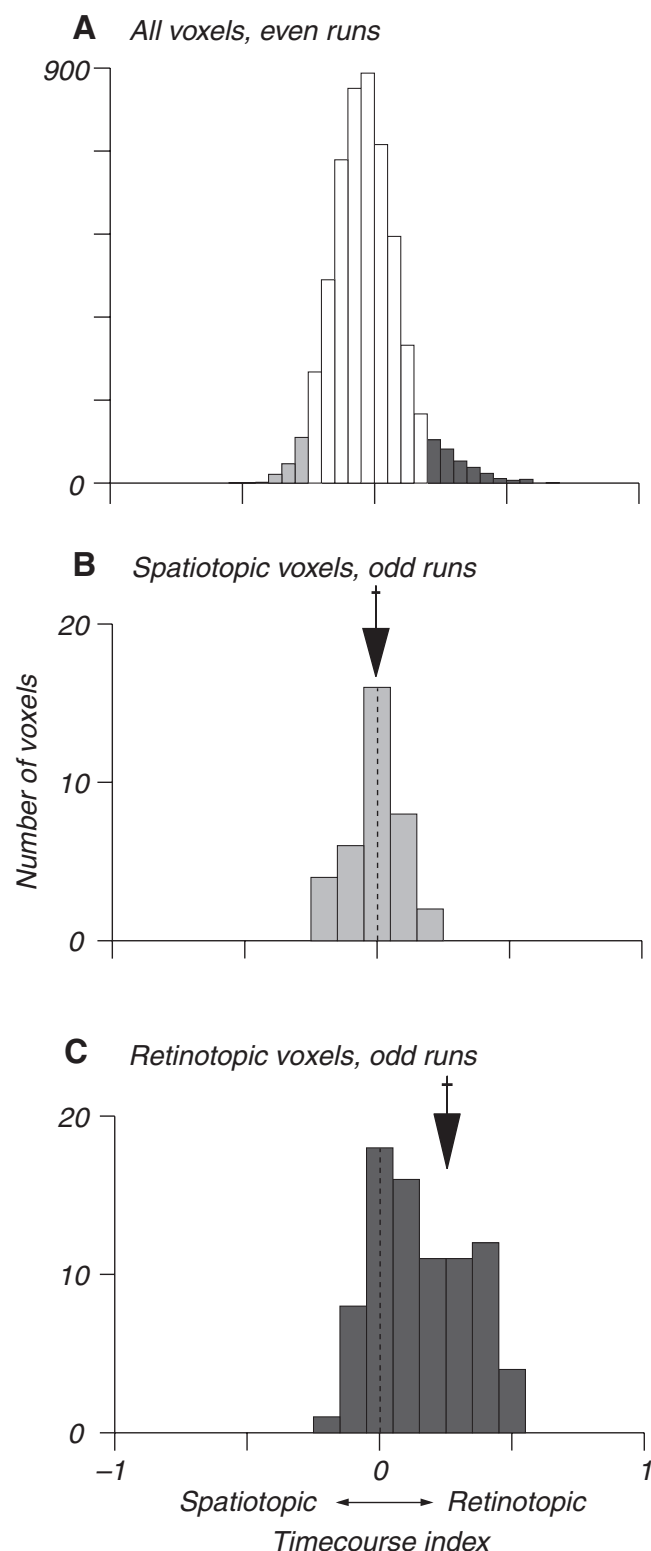


Figure 8. Cross-validation analysis shows that retinotopic, but not spatiotopic, responses were reliable. **A**, Distribution of time course indexes for all MT voxels from all subjects, for one-half of the data. Dark gray corresponds to a subset of voxels classified as retinotopic with a reference frame index value >0.2 , and light gray corresponds to voxels with an index less than -0.2 , suggesting that they might be spatiotopic. **B**, Distribution of time course indexes for the second half of the data, from voxels classified as spatiotopic (light gray in **A**) in the first half of the data. **C**, Distribution of time course indexes for the second half of the data, from voxels classified as retinotopic (dark gray in **A**) in the first half of the data. The arrows mark the means of the distributions in **B** and **C**.

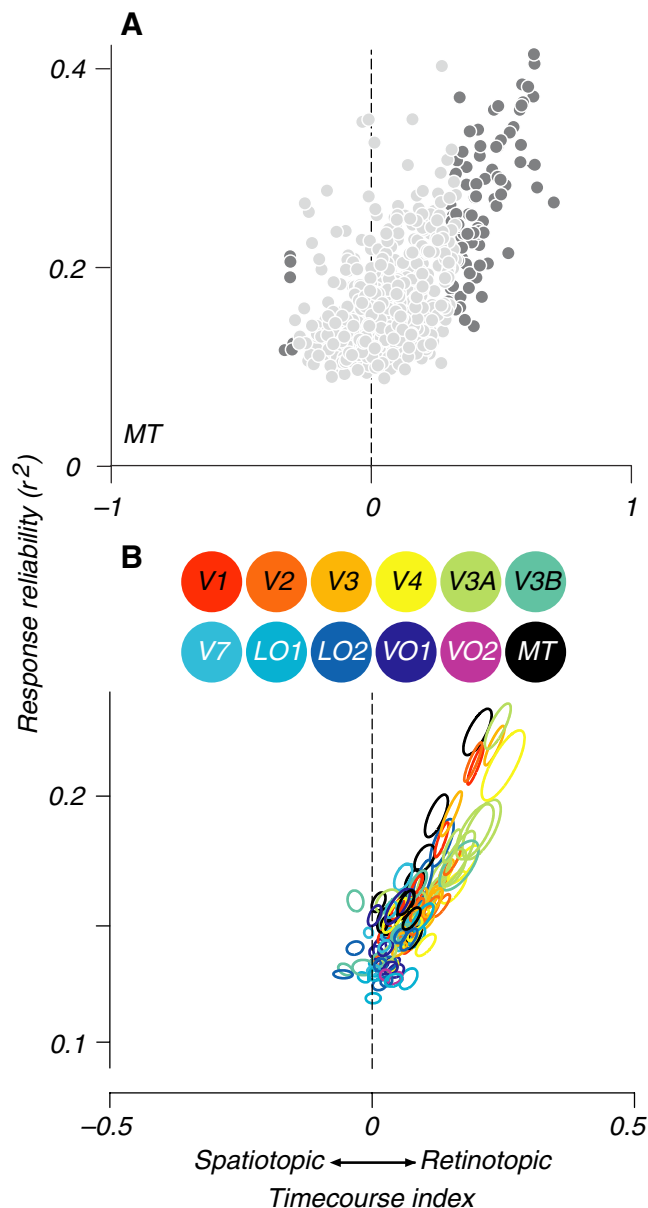


Figure 9. More reliable visual responses were more clearly retinotopic. **A**, Response reliability (r^2 value) plotted against reference frame index for each voxel within MT, combined across all subjects. Each data point represents a single voxel. Dark gray points represent reference frame index values that were significantly different from zero ($p < 0.01$, permutation statistical test). **B**, Summary of plots like that in **A** for all visual areas. Note change of scale from **A**. Each ellipse in **B** represents the mean and SE ellipse of the reliability versus reference frame, for one visual area in one subject. Ellipses are colored according to visual area (see legend above the graph).

responses to two visual motion localizers, not based on visuotopic mapping. The area that they called MT was small (0.6 cm^3 of cortex), less than one-quarter the size of MT in our study (median of 2.7 cm^3) (Fig. 2). It is impossible to determine the precise correspondence between MT and the area used by d'Avossa et al., because there is considerable variability across individuals in the locations of visual areas, including V1 and MT (Andrews et al., 1997; Amunts et al., 2000). Hence, anatomical references (e.g., Talairach coordinates) are not reliable indicators of visual areas (Saxe et al., 2006). Perhaps the area reported by d'Avossa et al. was a small anterior portion of what we call MT+ (which includes neighboring motion-sensitive areas such as

MST), and thus may not have included visuotopic MT at all. However, this explanation is unlikely because our explicit search for spatiotopic voxels in the vicinity of MT was unsuccessful. Second, our analysis took explicit account of the robustness of visual responses. Response time courses that are not robust cannot be used to determine the reference frame. Had d'Avossa et al. used an area outside visuotopic MT, without explicitly evaluating the robustness of visual responses, they may have been especially susceptible to misclassifying spurious responses as spatiotopic. Response amplitude estimates, even if they account for a significant portion of the variance, may not be faithful abstractions of the response time courses if the model they are based on does not provide a good fit to the data.

Can responses attributable to spatiotopic redirection of attention be confounded with a spatiotopic reference frame? MT activity can be modulated by task demands and attentional cues (Corbetta et al., 1990; Treue and Maunsell, 1996; Beauchamp et al., 1997; O'Craven et al., 1997; Gandhi et al., 1999; Huk and Heeger, 2000) and topographic activity can be recorded in human occipital regions by directing spatial attention (Brefczynski and DeYoe, 1999; Silver et al., 2005) even in the absence of a visual stimulus (Kastner et al., 1999; Ress et al., 2000; Serences and Boynton, 2007). If subjects directed their attention to a spatiotopic location regardless of fixation, for example the target on the left of the screen, responses would have been larger for the attended location for all eye positions. However, this possible attentional confound does not explain the results of d'Avossa et al. for two reasons. First, spatiotopic responses were reported both when subjects were instructed to attend to the targets and when their attention was not controlled at all. That is, attention was not a critical factor in the results reported by d'Avossa et al. Second and more decisively, attentional modulation can only explain the pattern of results seen in one hemisphere at a time. d'Avossa et al. reported that the left MT responded to the two targets on the right, regardless of eye position, whereas the right MT responded to the two targets on the left. If subjects directed their attention spatiotopically to the leftmost target, then spatiotopic responses in right MT might be explained, but the left MT would have also responded to the attended target on the left, and not to the targets on the right as was reported by d'Avossa et al.

Our results are also consistent with a large body of accumulated evidence that monkey and human area MT are functionally corresponding areas. Neurons in monkey MT form a homogeneous population readily identifiable by a distinct and robust selectivity for the direction of visual motion (Zeki, 1974; Maunsell and Van Essen, 1983). A region in lateral occipital cortex in humans, initially identified by its high sensitivity to visual motion (Zeki et al., 1991; Tootell et al., 1995a), exhibits many anatomical and physiological properties that are hallmarks of MT in monkeys. fMRI responses in this region show some direction selectivity (Huk et al., 2001; Kamitani and Tong, 2006), respond monotonically to increases in motion coherence (Rees et al., 2000), adapt selectively to directional motion (Tootell et al., 1995b; Huk et al., 2001), and are selective for pattern- as well as component-motion (Huk and Heeger, 2002). This region, dubbed MT+ or hMT, most likely corresponds to a combination of monkey areas MT and MST, and by using functional and visuotopic mapping criteria, can be subdivided into regions analogous to monkey MT and MST (Huk et al., 2002; Smith et al., 2006). The response characteristics of the subdivided human MT suggest that this region in humans is functionally equivalent to area MT in monkeys. Our results extend this functional equivalence by showing human MT, like monkey MT (Krekelberg et al., 2003), represents stimuli in a retinotopic reference frame. Indeed, our findings

demonstrate that retinotopic reference frames are a fundamental property that is shared not just between functionally equivalent areas in the human and monkey, but among all routinely identifiable occipital visual areas.

References

- Amunts K, Malikovic A, Mohlberg H, Schormann T, Zilles K (2000) Brodmann's areas 17 and 18 brought into stereotaxic space—where and how variable? *NeuroImage* 11:66–84.
- Andersen RA, Mountcastle VB (1983) The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *J Neurosci* 3:532–548.
- Andersen RA, Asanuma C, Cowan WM (1985a) Callosal and prefrontal associational projecting cell populations in area 7A of the macaque monkey: a study using retrogradely transported fluorescent dyes. *J Comp Neurol* 232:443–455.
- Andersen RA, Essick GK, Siegel RM (1985b) Encoding of spatial location by posterior parietal neurons. *Science* 230:456–458.
- Andrews TJ, Halpern SD, Purves D (1997) Correlated size variations in human visual cortex, lateral geniculate nucleus, and optic tract. *J Neurosci* 17:2859–2868.
- Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR (2005) Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci* 8:941–949.
- Baker JT, Donoghue JP, Sanes JN (1999) Gaze direction modulates finger movement activation patterns in human cerebral cortex. *J Neurosci* 19:10044–10052.
- Beauchamp MS, Cox RW, DeYoe EA (1997) Graded effects of spatial and featural attention on human area MT and associated motion processing areas. *J Neurophysiol* 78:516–520.
- Boynton GM, Engel SA, Glover GH, Heeger DJ (1996) Linear systems analysis of functional magnetic resonance imaging in human V1. *J Neurosci* 16:4207–4221.
- Brainard DH (1997) The Psychophysics Toolbox. *Spat Vis* 10:433–436.
- Brefczynski JA, DeYoe EA (1999) A physiological correlate of the “spotlight” of visual attention. *Nat Neurosci* 2:370–374.
- Bremmer F, Ilg UJ, Thiele A, Distler C, Hoffmann KP (1997) Eye position effects in monkey cortex. I. Visual and pursuit-related activity in extrastriate areas MT and MST. *J Neurophysiol* 77:944–961.
- Brewer AA, Liu J, Wade AR, Wandell BA (2005) Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nat Neurosci* 8:1102–1109.
- Cohen YE, Andersen RA (2002) A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci* 3:553–562.
- Colby CL, Duhamel JR, Goldberg ME (1993) Ventral intraparietal area of the macaque: anatomical location and visual response properties. *J Neurophysiol* 69:902–914.
- Corbetta M, Miezin FM, Dobmeyer S, Shulman GL, Petersen SE (1990) Attentional modulation of neural processing of shape, color, and velocity in humans. *Science* 248:1556–1559.
- Dale AM (1999) Optimal experimental design for event-related fMRI. *Hum Brain Mapp* 8:109–114.
- d'Avossa G, Tosetti M, Crespi S, Biagi L, Burr DC, Morrone MC (2007) Spatiotopic selectivity of BOLD responses to visual motion in human area MT. *Nat Neurosci* 10:249–255.
- Dean HL, Platt ML (2006) Allocentric spatial referencing of neuronal activity in macaque posterior cingulate cortex. *J Neurosci* 26:1117–1127.
- Desimone R, Ungerleider LG (1986) Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol* 248:164–189.
- DeSouza JF, Dukelow SP, Vilis T (2002) Eye position signals modulate early dorsal and ventral visual areas. *Cereb Cortex* 12:991–997.
- Duhamel JR, Colby CL, Goldberg ME (1992) The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255:90–92.
- Duhamel JR, Bremmer F, BenHamed S, Graf W (1997) Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389:845–848.
- Duhamel JR, Colby CL, Goldberg ME (1998) Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J Neurophysiol* 79:126–136.
- Engel SA, Rumelhart DE, Wandell BA, Lee AT, Glover GH, Chichilnisky EJ, Shadlen MN (1994) fMRI of human visual cortex. *Nature* 369:525.
- Engel SA, Glover GH, Wandell BA (1997) Retinotopic organization in hu-

- man visual cortex and the spatial precision of functional MRI. *Cereb Cortex* 7:181–192.
- Galletti C, Battaglini PP, Fattori P (1995) Eye position influence on the parieto-occipital area PO (V6) of the macaque monkey. *Eur J Neurosci* 7:2486–2501.
- Gandhi SP, Heeger DJ, Boynton GM (1999) Spatial attention affects brain activity in human primary visual cortex. *Proc Natl Acad Sci USA* 96:3314–3319.
- Gardner JL, Sun P, Waggoner RA, Ueno K, Tanaka K, Cheng K (2005) Contrast adaptation and representation in human early visual cortex. *Neuron* 47:607–620.
- Goldberg ME, Bruce CJ (1990) Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J Neurophysiol* 64:489–508.
- Graziano MS, Yap GS, Gross CG (1994) Coding of visual space by premotor neurons. *Science* 266:1054–1057.
- Groh JM, Trause AS, Underhill AM, Clark KR, Inati S (2001) Eye position influences auditory responses in primate inferior colliculus. *Neuron* 29:509–518.
- Hansen KA, Kay KN, Gallant JL (2007) Topographic organization in and near human visual area V4. *J Neurosci* 27:11896–11911.
- Huk AC, Heeger DJ (2000) Task-related modulation of visual cortex. *J Neurophysiol* 83:3525–3536.
- Huk AC, Heeger DJ (2002) Pattern-motion responses in human visual cortex. *Nat Neurosci* 5:72–75.
- Huk AC, Ress D, Heeger DJ (2001) Neuronal basis of the motion aftereffect reconsidered. *Neuron* 32:161–172.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22:7195–7205.
- Jay MF, Sparks DL (1984) Auditory receptive fields in primate superior colliculus shift with changes in eye position. *Nature* 309:345–347.
- Kamitani Y, Tong F (2006) Decoding seen and attended motion directions from activity in the human visual cortex. *Curr Biol* 16:1096–1102.
- Kastner S, Pisk MA, De Weerd P, Desimone R, Ungerleider LG (1999) Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron* 22:751–761.
- Krekelberg B, Kubischik M, Hoffmann KP, Bremmer F (2003) Neural correlates of visual localization and perisaccadic mislocalization. *Neuron* 37:537–545.
- Larsson J (2001) Imaging vision: functional mapping of intermediate visual processes in man. PhD thesis, Karolinska Institute.
- Larsson J, Heeger DJ (2006) Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26:13128–13142.
- Maunsell JH, van Essen DC (1983) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3:2563–2586.
- Maunsell JH, Van Essen DC (1983) Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *J Neurophysiol* 49:1127–1147.
- McKyton A, Zohary E (2007) Beyond retinotopic mapping: the spatial representation of objects in the human lateral occipital complex. *Cereb Cortex* 17:1164–1172.
- Medendorp WP, Goltz HC, Vilis T, Crawford JD (2003) Gaze-centered updating of visual space in human parietal cortex. *J Neurosci* 23:6209–6214.
- Merriam EP, Genovesi CR, Colby CL (2003) Spatial updating in human parietal cortex. *Neuron* 39:361–373.
- Merriam EP, Genovesi CR, Colby CL (2007) Remapping in human visual cortex. *J Neurophysiol* 97:1738–1755.
- Nakamura K, Colby CL (2002) Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc Natl Acad Sci USA* 99:4026–4031.
- Nestares O, Heeger DJ (2000) Robust multiresolution alignment of MRI brain volumes. *Magn Reson Med* 43:705–715.
- O'Craven KM, Rosen BR, Kwong KK, Treisman A, Savoy RL (1997) Voluntary attention modulates fMRI activity in human MT-MST. *Neuron* 18:591–598.
- Ogawa S, Lee TM, Kay AR, Tank DW (1990) Brain magnetic resonance imaging with contrast dependent on blood oxygenation. *Proc Natl Acad Sci USA* 87:9868–9872.
- Olson CR, Gettner SN (1995) Object-centered direction selectivity in the macaque supplementary eye field. *Science* 269:985–988.
- Pelli DG (1997) The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* 10:437–442.
- Rees G, Friston K, Koch C (2000) A direct quantitative relationship between the functional properties of human and macaque V5. *Nat Neurosci* 3:716–723.
- Ress D, Backus BT, Heeger DJ (2000) Activity in primary visual cortex predicts performance in a visual detection task. *Nat Neurosci* 3:940–945.
- Saxe R, Brett M, Kanwisher N (2006) Divide and conquer: a defense of functional localizers. *NeuroImage* 30:1088–1099.
- Serences JT, Boynton GM (2007) Feature-based attentional modulations in the absence of direct visual stimulation. *Neuron* 55:301–312.
- Sereno MI, Huang RS (2006) A human parietal face area contains aligned head-centered visual and tactile maps. *Nat Neurosci* 9:1337–1343.
- Sereno MI, Dale AM, Reppas JB, Kwong KK, Belliveau JW, Brady TJ, Rosen BR, Tootell RB (1995) Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268:889–893.
- Siegel RM, Raffi M, Phinney RE, Turner JA, Jando G (2003) Functional architecture of eye position gain fields in visual association cortex of behaving monkey. *J Neurophysiol* 90:1279–1294.
- Silver MA, Ress D, Heeger DJ (2005) Topographic maps of visual spatial attention in human parietal cortex. *J Neurophysiol* 94:1358–1371.
- Smith AT, Wall MB, Williams AL, Singh KD (2006) Sensitivity to optic flow in human cortical areas MT and MST. *Eur J Neurosci* 23:561–569.
- Soechting JF, Flanders M (1992) Moving in three-dimensional space: frames of reference, vectors, and coordinate systems. *Annu Rev Neurosci* 15:167–191.
- Stricanne B, Andersen RA, Mazzoni P (1996) Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* 76:2071–2076.
- Swisher JD, Halko MA, Merabet LB, McMains SA, Somers DC (2007) Visual topography of human intraparietal sulcus. *J Neurosci* 27:5326–5337.
- Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci* 6:134–144.
- Tolias AS, Moore T, Smirnakis SM, Tehovnik EJ, Siapas AG, Schiller PH (2001) Eye movements modulate visual receptive fields of V4 neurons. *Neuron* 29:757–767.
- Tootell RB, Hadjikhani N (2001) Where is “dorsal V4” in human visual cortex? Retinotopic, topographic and functional evidence. *Cereb Cortex* 11:298–311.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW (1995a) Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15:3215–3230.
- Tootell RB, Reppas JB, Dale AM, Look RB, Sereno MI, Malach R, Brady TJ, Rosen BR (1995b) Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* 375:139–141.
- Treue S, Maunsell JH (1996) Attentional modulation of visual motion processing in cortical areas MT and MST. *Nature* 382:539–541.
- Umeno MM, Goldberg ME (1997) Spatial processing in the monkey frontal eye field. I. Predictive visual responses. *J Neurophysiol* 78:1373–1383.
- Walker MF, Fitzgibbon EJ, Goldberg ME (1995) Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *J Neurophysiol* 73:1988–2003.
- Wandell BA, Dumoulin SO, Brewer AA (2007) Visual field maps in human cortex. *Neuron* 56:366–383.
- Wetherill G, Levitt H (1965) Sequential estimation of points on a psychometric function. *Br J Math Stat Psychol* 18:1–10.
- Zeki S, Watson JD, Lueck CJ, Friston KJ, Kennard C, Frackowiak RS (1991) A direct demonstration of functional specialization in human visual cortex. *J Neurosci* 11:641–649.
- Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol (Lond)* 236:549–573.