

Eye-centered visual receptive fields in the ventral intraparietal area

Xiaodong Chen,^{1,2} Gregory C. DeAngelis,³ and Dora E. Angelaki^{1,2}

¹Department of Neuroscience, Baylor College of Medicine, Houston, Texas; ²Department of Anatomy & Neurobiology, Washington University, St. Louis, Missouri; and ³Department of Brain and Cognitive Sciences, University of Rochester, Rochester, New York

Submitted 17 January 2014; accepted in final form 29 April 2014

Chen X, DeAngelis GC, Angelaki DE. Eye-centered visual receptive fields in the ventral intraparietal area. *J Neurophysiol* 112: 353–361, 2014. First published April 30, 2014; doi:10.1152/jn.00057.2014.—The ventral intraparietal area (VIP) processes multisensory visual, vestibular, tactile, and auditory signals in diverse reference frames. We recently reported that visual heading signals in VIP are represented in an approximately eye-centered reference frame when measured using large-field optic flow stimuli. No VIP neuron was found to have head-centered visual heading tuning, and only a small proportion of cells had reference frames that were intermediate between eye- and head-centered. In contrast, previous studies using moving bar stimuli have reported that visual receptive fields (RFs) in VIP are head-centered for a substantial proportion of neurons. To examine whether these differences in previous findings might be due to the neuronal property examined (heading tuning vs. RF measurements) or the type of visual stimulus used (full-field optic flow vs. a single moving bar), we have quantitatively mapped visual RFs of VIP neurons using a large-field, multipatch, random-dot motion stimulus. By varying eye position relative to the head, we tested whether visual RFs in VIP are represented in head- or eye-centered reference frames. We found that the vast majority of VIP neurons have eye-centered RFs with only a single neuron classified as head-centered and a small minority classified as intermediate between eye- and head-centered. Our findings suggest that the spatial reference frames of visual responses in VIP may depend on the visual stimulation conditions used to measure RFs and might also be influenced by how attention is allocated during stimulus presentation.

parietal cortex; VIP; reference frame; receptive field; eye centered

VISUAL SIGNALS ARE INTEGRATED with other sensory cues in many cortical areas to construct a more complete representation of space (Avillac et al. 2005; Cohen and Andersen 2002; Fetsch et al. 2007; Maier and Groh 2009; Pouget et al. 2002). The ventral intraparietal area (VIP) is a multimodal brain region receiving sensory inputs from visual, vestibular, auditory, and somatosensory systems (Avillac et al. 2005; Bremmer et al. 2002a,b; Chen et al. 2011a,b; Colby et al. 1993; Duhamel et al. 1998; Guipponi et al. 2013; Lewis and Van Essen 2000; Maciokas and Britten 2010; Schlack et al. 2002, 2005; Zhang and Britten 2010). These different sensory signals in VIP are represented in diverse spatial reference frames. Specifically, facial tactile receptive fields (RFs) are represented in a head-centered reference frame (Avillac et al. 2005), whereas visual and auditory RFs are organized in a continuum between eye- and head-centered coordinates (Schlack et al. 2005). Vestibular heading tuning in VIP is represented in a body-centered refer-

ence frame (Chen et al. 2013b), whereas visual heading tuning (based on optic flow) is coded mainly in an eye-centered reference frame (Chen et al. 2013c).

The eye-centered visual heading tuning of VIP neurons (Chen et al. 2013c) appears to be at odds with the previous reports of head-centered visual RFs from Duhamel et al. (1997) and Avillac et al. (2005). There are two main differences between these two groups of studies. The first difference involves the visual stimuli that were used: Chen et al. (2013c) used large-field optic flow (random dot) stimuli, whereas Duhamel et al. (1997) used a single bar stimulus that moved in the preferred direction of each neuron and thus stimulated one small patch of the visual field at a time. The second difference between studies is the neuronal response property that was examined: Chen et al. (2013c) measured visual heading tuning curves, whereas Duhamel et al. (1997) characterized visual RFs. Although there is no particular reason to believe that the spatial reference frames of visual heading tuning and visual RF locations should be linked, it is currently unclear whether the differences between previous studies are mainly due to the different response properties measured or the different stimuli used.

We have used a large-field stimulus consisting of multiple simultaneously presented random-dot patches, along with a reverse-correlation technique (Chen et al. 2008), to quantify the spatial and directional structure of visual RFs in VIP. We found that VIP visual RFs shifted systematically with eye position, largely consistent with an eye-centered reference frame. Although a small minority of cells was characterized by an intermediate representation, our results differ substantially from those of Duhamel et al. (1997) and Avillac et al. (2005), who reported that a substantial proportion of VIP neurons have head-centered RFs.

MATERIALS AND METHODS

Subjects and Experimental Protocols

Extracellular recordings were made from four hemispheres in two male rhesus monkeys (*Macaca mulatta*) weighing 7–10 kg. The monkeys were chronically implanted with a circular Delrin ring for head stabilization as well as two scleral search coils for measuring binocular eye position. Details have been described in previous publications (Chen et al. 2013a,b,c; Fetsch et al. 2007; Gu et al. 2006; Takahashi et al. 2007). All procedures were approved by the Institutional Animal Care and Use Committee at Washington University in St. Louis and were in accordance with National Institutes of Health guidelines.

During experiments, the monkey was seated comfortably in a primate chair secured on a motion platform. Visual stimuli, generated using the OpenGL graphics library and an OpenGL accelerator board [Quadro FX 3000G; PNY Technologies, Parsippany, NJ; see Gu et al. (2006) for details], were rear-projected (Christie Digital Mirage 2000;

Address for reprint requests and other correspondence: D. E. Angelaki, Dept. of Neuroscience, BCM295, Baylor College of Medicine, One Baylor Plaza, Rm. T504C, Houston, TX 77030 (e-mail: angelaki@cabernet.cns.bcm.edu).

Cyrus, CA) onto a tangent screen placed 30 cm in front of the monkey (subtending 90° of visual angle). The monkey chair and the tangent screen were covered on all sides with black matte material such that the monkey's field of view was restricted to the visual display. Image resolution was 1,280 × 1,024 pixels, and refresh rate was 60 Hz. Visual stimuli were viewed binocularly at zero disparity.

Visual RFs were measured using a multipatch, random-dot motion stimulus (Fig. 1) along with a reverse-correlation technique (Chen et al. 2008). The visual stimulus, which covered an area that subtended 80 × 80°, was divided into a virtual square grid of 64 (8 × 8) nonoverlapping subfields with each subfield subtending 10 × 10° of visual angle. In each video frame, each subfield contained a pattern of 0.15- × 0.15-cm yellow dots (density: 0.01/cm²) that were located at random positions (dots are shown as gray in Fig. 1). Across video frames, all dots within each subfield moved coherently in 1 of 8 possible directions (0, 45, 90, 135, 180, 225, 270, or 315°; 0°: rightward; 90°: upward). The directions of motion in every subfield were updated simultaneously and independently every 100 ms throughout each 2-s trial. The same speed of motion (40°/s) was applied to dots in all subfields because most VIP neurons prefer high speeds (Bremmer et al. 2002a; Yang et al. 2011).

In each trial, monkeys were required to fixate a target (0.2 × 0.2°) for 200 ms before stimulus onset and to maintain fixation throughout the whole trial (2 s) to receive a reward. In all experiments here, the head was fixed relative to the body, and only eye position relative to the head was varied. Thus head- and body-centered representations are indistinguishable in these experiments. The fixation target was presented at one of three possible positions spaced 20° apart (0° at the center of the screen, -20° to the left of center, or +20° to the right of center), and the location was chosen randomly in each trial. Thus three separate RF maps were obtained, corresponding to the three eye positions. Trials were aborted and data were discarded when a monkey's gaze deviated by >1° from the fixation target. Typically, about 60–100 trials (mean = 80) were necessary to obtain a reasonably smooth RF map for each fixation position (Chen et al. 2008).

Neural Recordings

For single-unit recordings, a plastic grid made from Delrin (3.5 × 5.5 × 0.5 cm), containing staggered rows of holes (0.8-mm spacing), was stereotaxically attached to the inside of the head-restraint ring using dental acrylic and was positioned to overlay VIP in both hemispheres. The patterns of white and gray matter, as well as neuronal response properties, were used to identify VIP as described previously (Chen et al. 2011a,b, 2013a,b,c). Recordings were made using tungsten microelectrodes (FHC, Bowdoinham, ME) that were inserted into the brain via transdural guide tubes. For each neuron encountered in an electrode penetration, we first explored the RF and tuning properties qualitatively by manually controlling the parameters of a flickering or moving random-dot stimulus and observing the instantaneous firing rate of the neuron in a graphical display. The RF mapping protocol, which allowed quantification of the RF for each eye position, was delivered after the preliminary mapping of visual response properties.

Data presented here were recorded from the same animals and locations as those studied by Chen et al. (2013c; see their Fig. 6 for recording locations; $n = 29$ from *monkey E*, $n = 57$ from *monkey Q*). Some VIP neurons ($n = 15$ from *monkey E*, $n = 36$ from *monkey Q*) were tested with both visual heading tuning [see Chen et al. (2013c) for details] and RF mapping protocols, whereas the remaining neurons were tested only with the RF mapping protocol. Data for visual heading tuning have been presented elsewhere (Chen et al. 2013c; see their Fig. 3), so we focus on the RF mapping data in this report. Results were similar for the two monkeys, thus data were pooled across monkeys for all histograms and population analyses.

Analysis of RF Maps

All data analyses were done in MATLAB (MathWorks, Natick, MA). To obtain a RF map for each neuron, an 80 × 80° area of the display screen was divided into 64 (8 by 8) equal-sized subfields, and a random sequence of directions of motion was presented in each subfield. Neural responses to the stimuli in each subfield were characterized by the strength of directional selectivity, which was measured using a reverse-correlation method (Chen et al. 2008). Specifically, the spike train was cross-correlated with the temporal sequence of motion directions to generate a direction-time response map for each subfield in the stimulus grid. If neuronal responses are coupled with the stimulus over a range of correlation delays (T) from 0 to 200 ms (in 1-ms steps), a pattern will emerge in the direction-time response map; otherwise, the profiles will show no structure. These maps, therefore, reveal the direction tuning for each subfield that falls within the RF of the neuron. Note that this particular method will only reveal portions of the RF that have directional selectivity. In contrast, portions of RF that are nondirection-selective will not be registered by this method (Chen et al. 2008).

For each subfield, we identified the peak correlation delay (T_{peak}) as the value of T at which the variance among direction tuning curves in all 64 subfields reached a maximum. A horizontal cross-section through the response maps at T_{peak} yields a direction tuning curve for each subfield in the stimulus grid (Chen et al. 2008). The strength of direction tuning was estimated by computing the vector sum of the normalized responses to the 8 directions of motion. The larger the vector sum value, the stronger and/or narrower the direction tuning. A confidence interval (CI) was calculated for each vector sum value using a resampling method (Chen et al. 2008). Randomized direction tuning curves were generated from a range of negative (i.e., non-causal) correlation delays from 0 to -200 ms (in 1-ms steps) for each subfield, and a vector sum was calculated. These 200 vector sums formed a distribution reflecting the noise level in the measurements, from which a 95% CI was derived (percentile method). The direction tuning in a particular subfield was considered statistically significant if its corresponding vector sum lay outside of the 95% CI generated from the noncausal correlation delays. Only neurons for which the RF maps for all 3 eye positions had vector sum values significantly different from the noise level (for 5 sequential correlation delays) are included in the following analyses ($n = 62$, 26 from *monkey E*, 36 from *monkey Q*).

For neurons that met these criteria, the direction tuning curve for each subfield was fit with a von Mises function (Chen et al. 2008, 2013b,c; Fetsch et al. 2007) given by:

$$R(\theta) = A \cdot e^{\frac{-2[1 - \cos(\theta - \theta_p)]}{\sigma^2}} + r_b \quad (1)$$

where θ denotes heading, θ_p denotes the preferred heading, σ determines tuning width, A represents the response amplitude, and r_b denotes the baseline firing rate. The neuronal response strength in each subfield was estimated as the peak-to-trough response amplitude of the fitted von Mises curve. These amplitude values were then plotted as a function of the center location of each stimulus subfield, thus creating a two-dimensional (2-D) RF map.

Quantification of RF Shifts

To determine the spatial reference frame of visual RFs for VIP neurons, we used three different methods to quantify how RF maps shifted with eye position.

Gaussian function fits. The location and size of the RF were quantified by fitting the amplitude data with a 2-D Gaussian function (Chen et al. 2008):

$$f(x, y) = A \cdot e^{-\frac{\left(\frac{x-x_c}{\sigma_x}\right)^2 + \left(\frac{y-y_c}{\sigma_y}\right)^2}{2}} + b \quad (2)$$

where (x_c, y_c) are the coordinates of the peak of the spatial profile, σ_x and σ_y are the tuning widths along the horizontal (x) and vertical (y) dimensions, respectively, and b is the baseline response level. The goodness of fit was quantified by calculating the correlation coefficient between the original RF and the fitted 2-D Gaussian function as indicated by r^2 values. For the majority of fits (75.8%), r^2 values were >0.6 (Fig. 5A). Shifts in RF location with eye position could be assessed as changes in the value of x_c for RF maps measured at the three different horizontal eye positions. In addition, response gain field modulations by eye position were measured by the amplitude ratios, A_{L20}/A_0 (which quantified the gain change between left and center eye positions) and A_{R20}/A_0 (gain change between right and center positions). Only neurons for which all three RFs were well-fit separately by Eq. 2 ($r^2 > 0.6$, all 6 fit parameters free) were included in the analyses of gain fields and RF shifts ($n = 42$, 17 from *monkey E*, 25 from *monkey Q*).

Displacement index. The spatial reference frame of visual RFs was also assessed nonparametrically by measuring RF shifts using a displacement index (DI):

$$DI_{i,j} = \frac{k_{x,y}^{\max\{\text{cov}[R_i(x,y), R_j(x+k_x, y+k_y)]\}}}{P_i - P_j} \quad (3)$$

Here, k_x and k_y (in degrees) are the relative displacements of the two RFs (denoted R_i and R_j) along the x - and y -axes, respectively, and the superscript above k refers to the maximum covariance between the two RFs, which were linearly interpolated to 1° resolution (using the MATLAB function `interp2`) and systematically displaced relative to one another along the x - and y -axes. The denominator represents the difference between the two eye-fixation positions (P_i and P_j) at which the RFs were measured. Note that only the horizontal component of DI was used for this analysis because eye position only varied horizontally. If a RF shifts by an amount equal to the change in eye position, then $DI = 1$, and the RF is considered to be eye-centered. If no shift in RF position occurs with changes in eye position, then $DI = 0$, and the RF is head-centered. Only cells with good RF fits for at least two eye positions (r^2 of 2-D Gaussian fit >0.6 , Eq. 2) were included in the DI analyses. The number of RF pairs (i.e., number of DIs) that passed all criteria to be included in this analysis was $n = 139$ (59 pairs from 24 neurons in *monkey E*, 80 pairs from 33 neurons in *monkey Q*). Although T_{peak} could be different for each eye position, differences in T_{peak} across eye positions were generally quite small. Moreover, there was no significant correlation at the population level between DI and the difference in T_{peak} between eye positions ($P = 0.21$, Wilcoxon signed-rank test). Results were very similar if a single value of T_{peak} (from the RF map with the maximum peak response) was used across all three eye positions.

A CI was computed for each DI value using a bootstrap method. Bootstrapped RFs were generated by resampling (with replacement) the data for each motion direction, and then a DI value was computed for each pair of bootstrap RFs. This was repeated 1,000 times to produce a distribution of bootstrap DI values from which a 95% CI was derived (percentile method). A DI value was considered significantly different from a particular value (0 and/or 1) if its 95% CI did not include that value. A DI was classified as eye-centered if the CI did not include 0 but included 1. A DI was classified as head-centered if the CI did not include 1 but included 0. Finally, a DI was classified as intermediate if the CI was contained in the interval between 0 and 1 but did not include 0 or 1. All other cases were designated as unclassified.

DI values from RF data are also compared with DI values from visual heading tuning curves as long as the data corresponding to both DI values passed the following criteria for significant spatial structure. For the former (RF DIs), the requirement was that RF maps for both eye positions should be well-fit (r^2 of 2-D Gaussian fit > 0.6). For the

latter (heading tuning DIs), significance was judged as follows [see Chen et al. (2013c) for details]. Peristimulus time histograms were constructed for each heading and each eye-fixation position. Heading tuning curves were then constructed by plotting firing rate, computed in a 400-ms window centered on the “peak time” of the neuron (Chen et al. 2010), as a function of heading. The peak time was defined as the center of the 400-ms window for which the neuronal response reached its maximum across all stimulus conditions. To identify the peak time, firing rates were computed in many different 400-ms time windows spanning the range of the data in 25-ms steps. For each 400-ms window, a one-way ANOVA (response by heading direction) was performed. Heading tuning was considered statistically significant if the one-way ANOVA passed the significance test ($P < 0.05$) for five contiguous time points centered on the peak time. Only DIs computed from significant heading tuning curves were included. The number of cases for which both the RF maps and the heading tuning data passed the criteria for computing DI values was $n = 48$ (18 from *monkey E*, 30 from *monkey Q*).

Fitting RFs with eye- and head-centered models. To determine whether the set of RFs measured at all three eye positions was most consistent with an eye- or head-centered representation, the three RF profiles for each neuron were fit simultaneously with a set of functions having the form of Eq. 2. In the eye-centered model, A , b , y_c , σ_x , and σ_y were free parameters for each eye position, but x_c was constrained to shift by exactly the amount of the eye position change (i.e., x_c for straight ahead, $x_c + 20^\circ$ for left fixation, and $x_c - 20^\circ$ for right fixation). In the head-centered model, x_c was constrained to have the same value for all 3 eye positions (no shift), but the other 5 parameters were free to vary with eye position. Thus the total number of free parameters for each model was 16 (5 free parameters times 3 fixation positions, plus x_c).

For each fit, the correlation between the best-fitting function and the data was computed to measure goodness of fit. To remove the influence of correlations between the models themselves, partial correlation coefficients were computed by the MATLAB function `partialcorr` and subsequently normalized using Fisher r -to- z transform (Angelaki et al. 2004; Chen et al. 2013b,c; Fetsch et al. 2007; Smith et al. 2005) such that z -scores from the eye- and head-centered models could be compared in a scatterplot. If the z -score for 1 model was >2.326 and exceeded the z -score for the other model by ≥ 2.326 (equivalent to a P value of 0.01), that model was considered a significantly better fit to the data than the alternative model (Chen et al. 2013b,c; Fetsch et al. 2007). Only neurons for which all 3 RF maps were well-fit separately by Eq. 2 ($r^2 > 0.6$, all 6 parameters free) were included in this z -score analysis ($n = 42$, 17 from *monkey E*, 25 from *monkey Q*).

RESULTS

To examine whether VIP RFs are represented in eye- or head-centered reference frames, RF maps were obtained for three different eye-fixation positions (-20° left, 0° center, and 20° right) while keeping the head and body fixed relative to the world. As illustrated in Fig. 1, we used a multipatch, random-dot visual motion stimulus to map RFs [see MATERIALS AND METHODS and Chen et al. (2008) for details]. Each of the 64 subfields (8×8 grid) was $10 \times 10^\circ$ in size and contained a random-dot pattern moving coherently in 1 of 8 directions (0° : rightward; 90° : upward). Every 100 ms, the direction of motion was chosen randomly for each of the 64 subfields such that each RF subregion was probed repeatedly with all 8 motion directions over time. The resulting spike train was cross-correlated with the stimulus sequence, resulting in a distinct direction-time map (Fig. 2, left column) for each individual subfield and each eye position. If the portion of the RF of the neuron overlying a particular stimulus subfield produces a direction-selective visual response, then that subfield will show

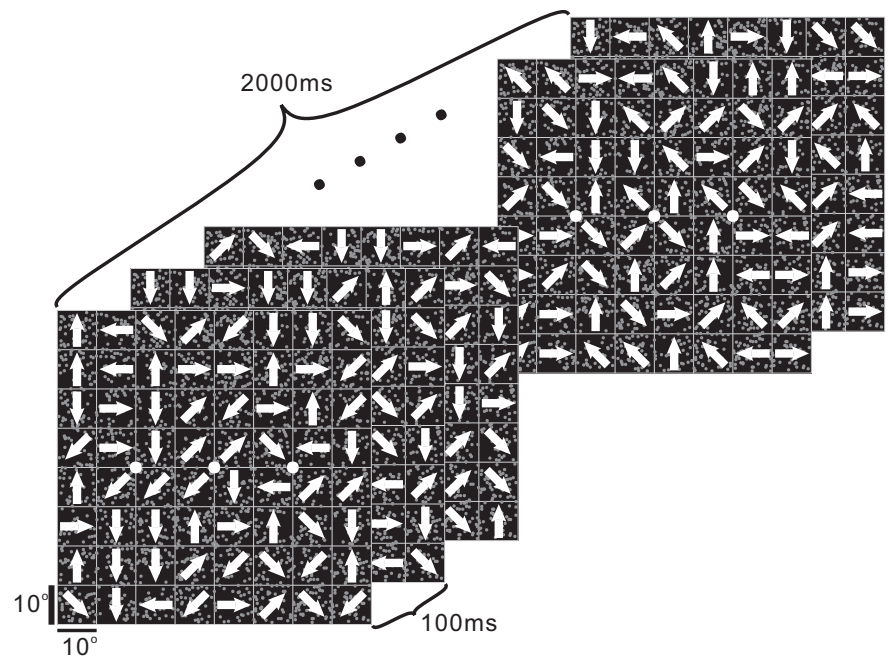


Fig. 1. Visual stimulus and protocol for receptive field (RF) mapping. The display screen ($80 \times 80^\circ$) was virtually divided into an 8×8 grid of subfields (each $10 \times 10^\circ$ in size). Random-dot motion stimuli were presented in each subfield simultaneously. At each location, random dots moved coherently (white arrows) in 1 of 8 possible directions (0, 45, 90, 135, 180, 225, 270, or 315°) at a speed of $40^\circ/\text{s}$. The direction of motion in each subfield was updated every 100 ms. In each trial, the monkey fixated at 1 of 3 possible locations spaced 20° apart horizontally (large white dots).

structure in the direction-time map. If no directional response is elicited by the stimulus in a particular subfield, the direction-time map will be unstructured.

For the example neuron shown in Fig. 2, a few subfields show clear directional responses, with a preference for $\sim 225^\circ$ (downward and leftward) motion. A horizontal cross-section through the direction-time map of each subfield at T_{peak} yields a direction tuning curve (Fig. 2, right column) for each subfield. These direction tuning curves were then fitted with von Mises functions (Eq. 1), and the amplitude difference between the peak and trough of the fitted tuning curve was taken as a measure of response strength for each subfield location. A spatial RF profile was then constructed by plotting this response strength as a function of the center location of each subfield as illustrated for the example VIP neuron in Fig. 3A. The three constructed RFs, one for each eye position, were then fitted with 2-D Gaussian functions (Eq. 2) as shown in Fig. 3B. For this example neuron, response amplitude varies substantially with eye position, indicative of a gain field with peak response increasing systematically from -20° (left) to 0° and to 20° (right). Importantly, the RF location relative to the fixation target (white x) remained relatively constant, whereas RF location relative to the screen clearly changed. This pattern indicates an eye-centered RF.

To quantify how the location of visual RFs changed with eye position, a DI was computed to quantify the shift of each pair of RFs relative to the change in eye position (Avillac et al. 2005; Fetsch et al. 2007). This method used the actual RF maps (not the Gaussian fits), which were interpolated and then shifted systematically relative to each other to find the shift that maximizes the cross-covariance between each pair of RFs (see MATERIALS AND METHODS for details). This technique takes into account the entire RF profile rather than just one parameter such as the peak. In addition, DI is also robust to variations in the gain or width of the RFs.

For the example neuron shown in Fig. 3A, $\text{DI} = 0.75$ for RFs measured at eye positions of -20 and 0° , $\text{DI} = 1.10$ for eye

positions of 0 and 20° , and $\text{DI} = 0.93$ for positions 20 and -20° . These values are broadly consistent with an eye-centered representation. Figure 4A shows the distribution of DI values obtained from two monkeys. DI values clustered near 1 with a mean DI of 0.92 ± 0.05 SE, which was significantly different from 0 ($P < 0.0001$, t -test) and significantly different from 1 ($P = 0.002$, t -test). At the single-neuron level, 105/139 DI values were classified as eye-centered, 19/139 DI values were classified as intermediate, and only 1/139 DI values were classified as head-centered. Thus the RFs of VIP neurons were generally eye-centered but with a small minority of RFs being significantly intermediate between head- and eye-centered.

The data of Fig. 4A show a substantial range of variation in DI values. This raises the question of whether this variability is mainly neural in origin or largely due to noise in our measurements and errors induced by our analysis procedures. In our experiment, eye position only varied horizontally. Thus we can examine the vertical shifts of RF location across fixation positions as a means to assess the contributions of noise and errors in analysis. Figure 4B shows the vertical shift in RF position plotted against the horizontal shift. Vertical shifts are generally very small, with a mean value of $0.018^\circ \pm 0.02$ SE that is not significantly different from 0 ($P = 0.062$, Wilcoxon signed-rank test) and is significantly smaller than the average horizontal shift value of $25.24^\circ \pm 2.01$ SE ($P < 0.0001$, Wilcoxon signed-rank test). The variance of the vertical shifts ($4.74^{\circ 2}$) is also highly significantly smaller than the variance of the horizontal shifts ($144.25^{\circ 2}$; Levene test, $P < 0.0001$). This suggests strongly that the range of DI values in Fig. 4A reflects real variation across neurons, not simply noise in the data or errors in our analysis.

The result of Fig. 4A appears to differ substantially from those of Duhamel et al. (1997) and Avillac et al. (2005), who reported that a substantial fraction of VIP neurons have head-centered RFs. Those previous studies constructed RF maps from raw neuronal responses (averaged spike counts) to a single oriented bar moving in the preferred direction of the

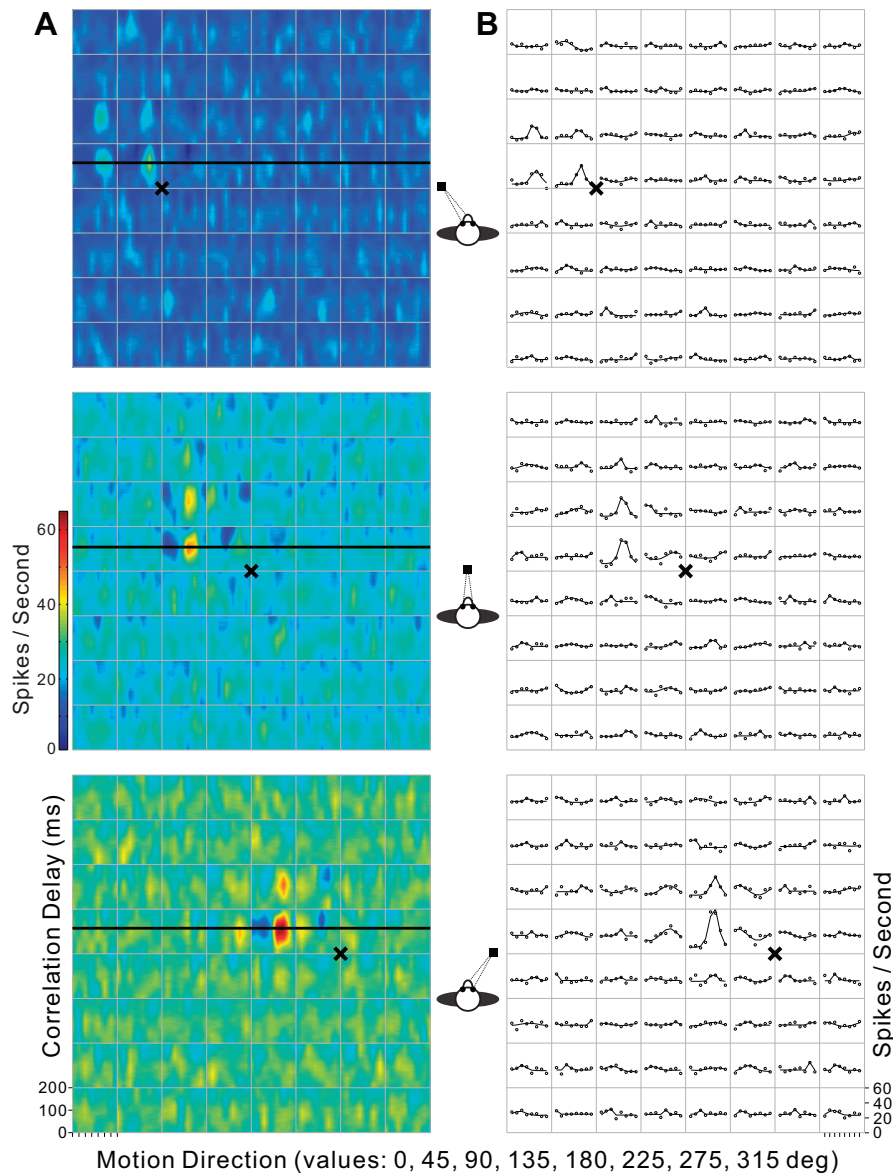


Fig. 2. RF profiles constructed by reverse correlation. A: direction-time RF maps are shown for each of the 8×8 subfields in the mapping grid. Color maps show neuronal responses as a function of motion direction (abscissa, 8 directions) and correlation delay (ordinate, 0–200 ms). The x-axis for each subplot ranges from 0 to 315° . B: direction tuning curves (circles) at the peak correlation delay (T_{peak} ; see MATERIALS AND METHODS for details) for each subfield along with the best-fitting von Mises functions (Eq. 1; solid curves). The 3 rows correspond to different eye-fixation locations (black crosses, -20 , 0 , and 20° from top to bottom). Black horizontal lines in A show the values of T_{peak} (110, 108, and 115 ms from top to bottom) for each of the 3 fixation locations. deg, Degrees.

neuron, whereas we used the peak-to-trough amplitude of the direction tuning curve as a measure of response strength for each RF subregion. To eliminate a potential difference in analysis method, we also quantified response strength for each subfield as the cumulative spike count at the preferred direction of the neuron, which was defined as the direction having the maximum summed response across all three eye positions. Using this alternative approach, we again found that DI values clustered near 1 with a mean DI value of 1.00 ± 0.05 SE, which was significantly different from 0 ($P < 0.0001$, t -test), not significantly different from 1 ($P = 0.53$, t -test), and not significantly different from DI values based on the original approach ($P = 0.08$, Wilcoxon signed-rank test). Moreover, DI values obtained using the two methods were well-correlated, with a slope that was not significantly different from unity [type II regression; $r = 0.56$, $P < 0.0001$, slope = 1.15, 95% CI = (0.87, 1.64)]. The average horizontal and vertical sizes of RFs (defined by the parameters of σ_x and σ_y in Eq. 2) were $\sim 10\%$ smaller using the alternative method of computing response strength, and this

difference was significant for vertical size ($P = 0.008$, Wilcoxon signed-rank test) but not for horizontal size ($P = 0.10$). Based on these observations, our approach of measuring response strength as the amplitude of the direction tuning curve for each RF subfield does not appear to account for the difference between our findings and those of Duhamel et al. (1997) and Avillac et al. (2005).

For a subset of neurons, we measured both RF maps and visual heading tuning curves using an optic flow stimulus [see Chen et al. (2013c) for details]. This allowed us to compare directly the reference frames of visual heading tuning and visual RFs for some VIP neurons. Median DI values computed from RF maps and heading tuning curves were not significantly different ($P = 0.09$, Wilcoxon signed-rank test). Moreover, they were correlated with a slope that was not significantly different from unity [type II regression; $r = 0.39$, $P = 0.006$, slope = 1.51, 95% CI = (0.82, 3.09); Fig. 4C]. Thus neurons with more eye-centered RFs also tended to have more eye-centered visual heading tuning curves, indicating that the spatial reference frames for these two response properties are related.

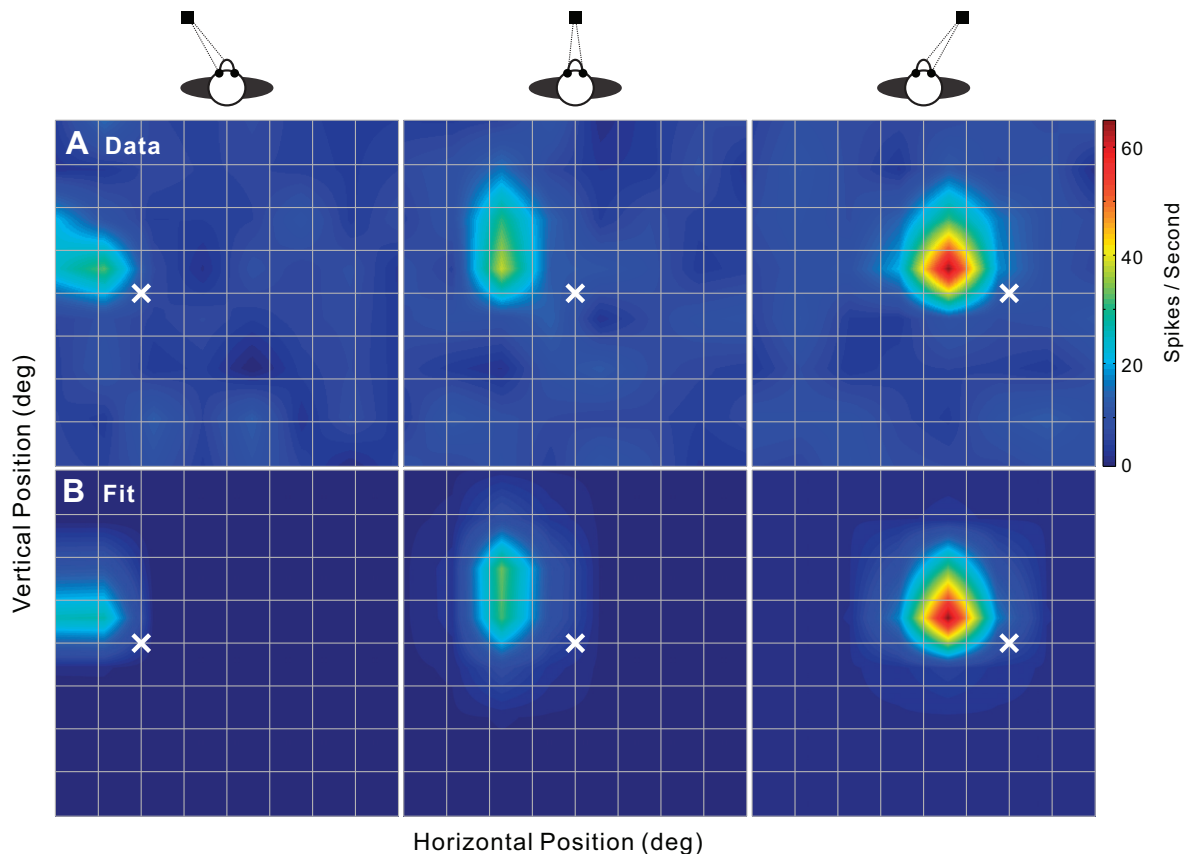


Fig. 3. Example RF profiles. *A*: RF maps for the example neuron of Fig. 2, shown for 3 fixation positions (white crosses, -20° , 0° , and 20° , from left to right). Firing rate is color-coded and plotted as a function of location in the stimulus grid. Each response value in these RF maps is computed from the amplitude of the corresponding direction tuning curve for that subfield (e.g., Fig. 2*B*). *B*: corresponding fits of a 2-dimensional (2-D) Gaussian function (Eq. 2) to the data of *A*.

To characterize better the effects of eye position variations on RF locations and response amplitudes, we also performed a model-based analysis using 2-D Gaussian fits (see MATERIALS AND METHODS and Fig. 3*B*). Overall, 2-D Gaussian functions provided good fits to the RF profiles of VIP neurons, as illustrated by the distribution of r^2 values in Fig. 5*A*. For cells with 2-D Gaussian fits having $r^2 > 0.6$ for all three eye positions, the shift in RF location between left (-20°) and center (0°) eye positions ($x_{L20} - x_0$) was plotted against the shift in RF location between right (20°) and center eye positions ($x_{R20} - x_0$), as shown in Fig. 5*B*. Most data points lie much closer to the eye-centered reference frame than to the head-centered reference frame, broadly consistent with the DI analysis.

The 2-D Gaussian fits also allowed us to characterize the effects of eye position on overall response gains. A substantial proportion of VIP neurons showed robust gain fields, and the gain ratios for left and right eye positions (A_{L20}/A_0 and A_{R20}/A_0) were significantly correlated, as illustrated in Fig. 5*C* [type II regression on logarithms of gain ratios; $r = 0.32$, $P = 0.038$, slope = 0.61, 95% CI = (0.15, 1.52)]. This positive correlation indicates that monotonic gain fields were not common. Rather, more data points fall in the upper right and lower left quadrants, consistent with nonmonotonic gain fields.

The predominantly eye-centered representation of RFs in VIP was further supported using an additional model fitting analysis to assess whether the overall pattern of data from each neuron is more consistent with eye- or head-centered reference frames. In this analysis, RF maps measured at different eye

positions were fitted simultaneously with an eye- and a head-centered model (see MATERIALS AND METHODS for details). The goodness of fit of each model was measured by computing correlation coefficients between the fitted RF and the data, which were then converted into partial correlation coefficients and normalized using Fisher r -to- z transform to enable meaningful comparisons between models independent of the number of data points (Angelaki et al. 2004; Chen et al. 2013b,c; Smith et al. 2005). Z -scores from the two models are compared in the scatterplot of Fig. 6. The gray region marks the boundaries of CIs that distinguish between models. Data points in the white area below the gray region were significantly better fit by the eye-centered model ($P < 0.01$), whereas data points above the gray region were significantly better fit by the head-centered model ($P < 0.01$). Data points located within the gray region were unclassified. In our sample of VIP neurons, 69.1% (29/42) were classified as eye-centered, 2.4% (1/42) were classified as head-centered, and 28.6% (12/42) were unclassified (Fig. 6). The distribution of differences in z -scores between the eye- and head-centered models is shown as the diagonal histogram. The mean difference in z -scores was 5.55, which is significantly greater than 0 (t -test, $P < 0.001$) and reinforces the conclusion that RF locations are represented predominantly in an eye-centered reference frame.

DISCUSSION

We systematically tested the spatial reference frames of visual RFs in VIP by measuring how RF locations shifted with

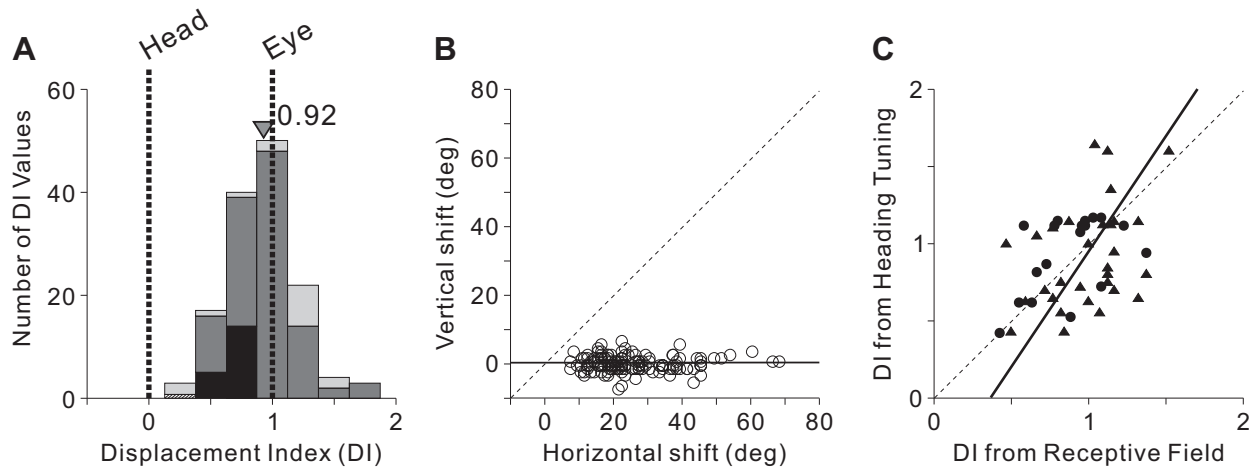


Fig. 4. Comparison of displacement indices (DIs) obtained from RF and heading tuning data. *A*: distribution of DI values from RF mapping data (shown only for pairs of fixation locations for which the 2-D Gaussian fits to the RF maps have $r^2 > 0.6$; *monkey E*: $n = 59$ pairs from 24 neurons; *monkey Q*: $n = 80$ pairs from 33 neurons). DI values of 0 and 1 indicate head- and eye-centered representations, respectively. Dark gray and hatched bars represent DI values that were classified as eye-centered ($n = 105$) and head-centered ($n = 1$), respectively. Black bars represent DI values that were classified as intermediate ($n = 19$), whereas light gray bars denote RF differences that were unclassified ($n = 14$). The arrowhead indicates the mean DI value. *B*: the vertical component of RF shift is plotted against the horizontal component of RF shift. Data are shown for the same set of pairs of fixation locations as in *A*. The black solid line represents a type II regression fit [$r = 0.04$, $P = 0.65$, slope = 0.007, 95% CI = (-0.03, 0.04)]; the dashed line indicates the unity-slope diagonal. *C*: DI values from RF data are plotted against DI values from heading tuning data (shown only for DI pairs that passed both heading tuning and RF significance criteria; see MATERIALS AND METHODS). Filled circles and triangles indicate data from *monkeys E* ($n = 18$) and *Q* ($n = 30$), respectively. The black solid line represents a type II regression fit; the dashed line indicates the unity-slope diagonal.

eye position using a reverse-correlation technique. Results from both empirical and model-based analyses show that the RFs of VIP neurons were generally represented in an eye-centered reference frame. A small proportion of neurons showed a reference frame intermediate between eye- and head-centered, similar to previous observations of intermediate reference frames in parietal cortex (Batista et al. 2007; Chang and Snyder 2010; Mullette-Gillman et al. 2005; Stricanne et al. 1996).

Chen et al. (2013c) previously showed that visual heading tuning in VIP was also represented in an eye-centered reference frame. Although the spatial reference frames of heading tuning and visual RF location need not be linked, similar observations were made for neurons in the dorsal medial superior temporal area, which are characterized by both eye-centered visual RFs (Lee et al. 2011) and eye-centered visual heading tuning (Fetsch et al. 2007). Moreover, for the subset of VIP neurons for which we measured the spatial reference frames of both visual RFs and visual heading tuning, the DI values were significantly correlated, indicating that cells with more (or less) eye-centered visual heading tuning also tend to have more (or less) eye-centered visual RFs.

Notably, our findings are somewhat different from those reported by Duhamel et al. (1997) and Avillac et al. (2005), who found that visual RFs in VIP showed a range of spatial reference frames from eye- to head-centered, with a substantial proportion of neurons exhibiting head-centered RFs. Duhamel et al. (1997) reported that head-centered RFs were more common when eye position was varied along the vertical (elevation) dimension than along the horizontal (azimuth) dimension. Since we restricted eye position changes to the horizontal meridian in the present study, this could account for a portion of the difference between studies. Nevertheless, even when considering only horizontal shifts in eye position, the data of Duhamel et al. (1997) and Avillac et al. (2005) appear

to be different from our findings. It is also worth noting that fMRI studies have revealed evidence consistent with head-centered visual RFs in parietal cortex (Sereno and Huang 2006).

One possible explanation for the difference between our findings and those of Duhamel and colleagues lies in the stimuli used. In the experiments of Duhamel et al. (1997), the stimulus was a white bar moving in the optimal direction within one subfield at a time. The motion lasted 100 ms, and, after 300 ms, the bar reappeared at a different location. In the present experiments, the stimulus used to characterize VIP RFs was a large-field, multipatch, random-dot stimulus in which motion was presented simultaneously at each of 64 subfield locations with independent random sequences of direction for each subfield. Thus key differences are that Duhamel et al. (1997) presented 1 moving bar at a time (whereas we stimulated all locations simultaneously), used only motion in the preferred direction (whereas direction varied randomly in our experiments), and used a higher speed (100 vs. our 40°/s). These differences were also shared by the experiments of Avillac et al. (2005). It is possible that presentation of many stimuli simultaneously engages lateral interactions or normalization mechanisms that might suppress head-centered inputs and leave the eye-centered inputs to dominate. At present, there is no direct evidence to support or refute this possibility.

Attention might also account for some of the difference between our findings and those of Duhamel et al. (1997). Although neither study controlled attention explicitly, attention may have been distributed differently due to the nature of the visual stimuli. When presented with a single moving bar stimulus at one subfield (Avillac et al. 2005; Duhamel et al. 1997), the monkey's covert attention may have been attracted away from the fixation point toward the moving bar. In contrast, with our full-field, multipatch, random-dot motion stimulus, no particular location on the screen was more salient than

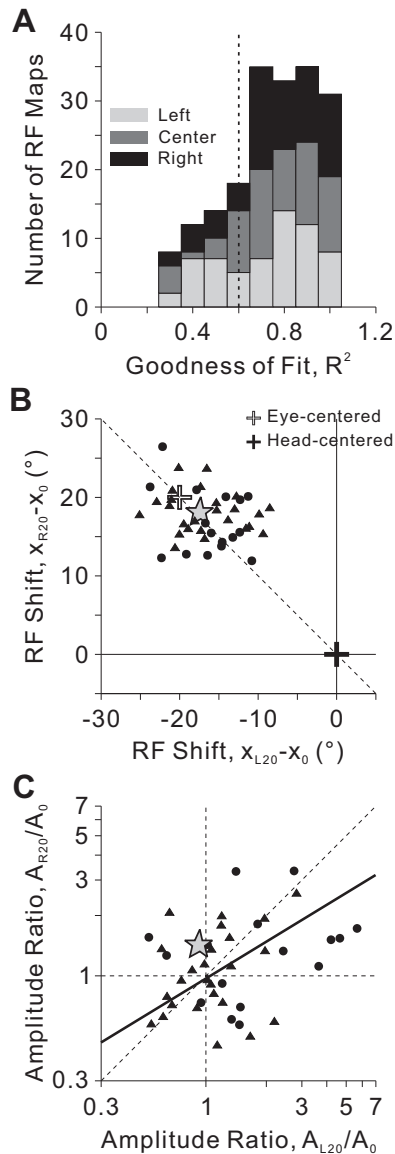


Fig. 5. Population summary of RF shifts and gain fields. *A*: distributions of r^2 values, which measure goodness of fit of the 2-D Gaussian function to individual RFs with significant structure (*monkey E*: $n = 78$; *monkey Q*: $n = 108$). Shades of gray correspond to the 3 eye positions. Vertical dashed line marks the value $r^2 = 0.6$ used as a criterion for inclusion in quantitative analyses. *B*: the difference in RF center location parameters (x_c in Eq. 2) between the left (-20°) and center (0°) eye-fixation positions ($x_{L20} - x_0$) is plotted against the difference in x_c between the right (20°) and center eye positions ($x_{R20} - x_0$). Unfilled and filled crosses represent eye- and head-centered reference frames, respectively. *C*: scatterplot of the amplitude ratios A_{L20}/A_0 (comparing left, -20° , and center, 0° , eye positions) and A_{R20}/A_0 (comparing right, 20° , and center positions). The solid line represents a type II regression fit (to the logarithm of the gain ratios). Dashed diagonal line and dashed vertical and horizontal unity amplitude ratio lines are also shown for comparison. Note that data are plotted on logarithmic scales for both axes. For *B* and *C*, filled circles and triangles denote data from *monkeys E* ($n = 17$) and *Q* ($n = 25$), respectively. The stars represent the example neuron of Figs. 2 and 3.

the others, and attention may have been directed to the fixation point a majority of the time. Alternatively, our animals may have distributed their attention over a broad region of the visual field. These potential differences in allocation of attention may have contributed to the different spatial reference frames found in the two studies, and future experiments should be designed to test this hypothesis.

Indeed, flexibility and diversity in reference frames according to both attentional and sensory context has been highlighted in recent neuroimaging studies. For example, head-centered representations in human MT+ were reported to arise only under conditions of spatially distributed attention and not while the subject performs a demanding task that requires allocation of attention near the fixation point (Burr and Morrone 2011). Accordingly, it was concluded that fMRI responses are retinotopic when attention is focused on the fixation point and spatiotopic when attention is allowed to be directed to the motion stimuli themselves (Burr and Morrone 2011). Although these conclusions were made for human MT+ rather than monkey VIP, it is possible that attention has analogous effects on spatial reference frames in both cases. Furthermore, fMRI activation in reach-coding areas has shown different reference frames when tested with visual vs. somatosensory stimuli. When targets were defined visually, the motor goal was encoded in gaze-centered coordinates; in contrast, when targets were defined by unseen proprioceptive cues, activity reflecting the motor goal was represented in body-centered coordinates (Bernier and Grafton 2010). In addition, the reference frames of auditory signals in primate superior colliculus have been reported to change dynamically from a hybrid frame of reference (intermediate between eye- and head-centered) to a predominantly eye-centered reference frame within the trial duration of a saccade task (Lee and Groh 2012). Collectively, these studies emphasize flexibility and diversity in spatial reference frame representations throughout the brain.

In summary, the current findings, combined with those of Chen et al. (2013c) and Duhamel and colleagues (Avillac et al. 2005; Duhamel et al. 1997), suggest that the reference frame representation of visual signals in VIP may depend on the experimental conditions under which visual RFs are measured, including the type of visual stimulus used and how attention is allocated. Moreover, these factors could be linked, as the nature of the visual stimulus may change how attention is allocated and, therefore, modulate spatial reference frames of

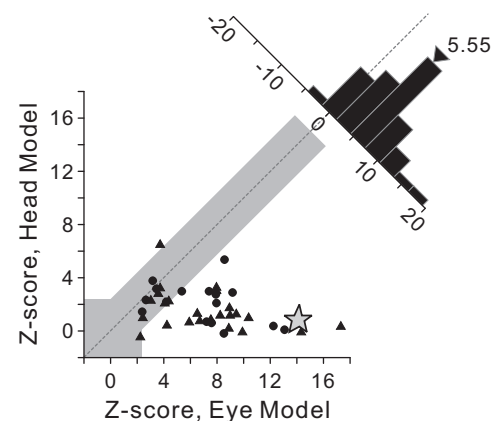


Fig. 6. Model-based classification of spatial reference frames for visual RFs in ventral intraparietal area. The scatterplot represents z-scored partial correlation coefficients between the data and either the eye- or head-centered model. Data points falling within the gray region cannot be classified, but data points falling into the areas above or below the gray region are classified as being best fit by the model represented on the ordinate or abscissa, respectively. Filled circles and triangles indicate data from the 2 monkeys (*monkey E*: $n = 17$; *monkey Q*: $n = 25$). The star represents the example neuron of Figs. 2 and 3. The diagonal histogram represents the distribution of differences between z-scores of the 2 models. Arrowhead indicates the mean value.

neural signals. Thus some of the flexibility and diversity of spatial reference frame representations in parietal cortex (Avillac et al. 2005; Bernier and Grafton 2010; Burr and Morrone 2011; Chen et al. 2013b,c; Duhamel et al. 1997; Mullette-Gillman et al. 2005, 2009; Schlack et al. 2005) may reflect dependencies on stimulus and task in addition to variations between areas.

ACKNOWLEDGMENTS

We thank Dr. E. Klier for helping with the writing as well as Amanda Turner and Jing Lin for excellent technical assistance.

GRANTS

This study was supported by National Eye Institute Grants R01-EY-017866 to D. E. Angelaki and R01-EY-016178 to G. C. DeAngelis.

DISCLOSURES

No conflicts of interest, financial or otherwise, are declared by the author(s).

AUTHOR CONTRIBUTIONS

X.C., G.C.D., and D.E.A. conception and design of research; X.C. performed experiments; X.C., G.C.D., and D.E.A. analyzed data; X.C., G.C.D., and D.E.A. interpreted results of experiments; X.C., G.C.D., and D.E.A. prepared figures; X.C. drafted manuscript; X.C., G.C.D., and D.E.A. edited and revised manuscript; X.C., G.C.D., and D.E.A. approved final version of manuscript.

REFERENCES

- Angelaki DE, Shaikh AG, Green AM, Dickman JD. Neurons compute internal models of the physical laws of motion. *Nature* 430: 560–564, 2004.
- Avillac M, Deneve S, Olivier E, Pouget A, Duhamel JR. Reference frames for representing visual and tactile locations in parietal cortex. *Nat Neurosci* 8: 941–949, 2005.
- Batista AP, Santhanam G, Yu BM, Ryu SI, Afshar A, Shenoy KV. Reference frames for reach planning in macaque dorsal premotor cortex. *J Neurophysiol* 98: 966–983, 2007.
- Bernier PM, Grafton ST. Human posterior parietal cortex flexibly determines reference frames for reaching based on sensory context. *Neuron* 68: 776–788, 2010.
- Bremmer F, Duhamel JR, Ben Hamed S, Graf W. Heading encoding in the macaque ventral intraparietal area (VIP). *Eur J Neurosci* 16: 1554–1568, 2002a.
- Bremmer F, Klam F, Duhamel JR, Ben Hamed S, Graf W. Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). *Eur J Neurosci* 16: 1569–1586, 2002b.
- Burr DC, Morrone MC. Spatiotopic coding and remapping in humans. *Philos Trans R Soc Lond B Biol Sci* 366: 504–515, 2011.
- Chang SW, Snyder LH. Idiosyncratic and systematic aspects of spatial representations in the macaque parietal cortex. *Proc Natl Acad Sci USA* 107: 7951–7956, 2010.
- Chen A, DeAngelis GC, Angelaki DE. A comparison of vestibular spatio-temporal tuning in macaque parieto-insular vestibular cortex, ventral intraparietal area, and medial superior temporal area. *J Neurosci* 31: 3082–3094, 2011a.
- Chen A, DeAngelis GC, Angelaki DE. Functional specializations of the ventral intraparietal area for multisensory heading discrimination. *J Neurosci* 33: 3567–3581, 2013a.
- Chen A, DeAngelis GC, Angelaki DE. Macaque parieto-insular vestibular cortex: responses to self-motion and optic flow. *J Neurosci* 30: 3022–3042, 2010.
- Chen A, DeAngelis GC, Angelaki DE. Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex. *J Neurosci* 31: 12036–12052, 2011b.
- Chen A, Gu Y, Takahashi K, Angelaki DE, DeAngelis GC. Clustering of self-motion selectivity and visual response properties in macaque area MSTd. *J Neurophysiol* 100: 2669–2683, 2008.
- Chen X, DeAngelis GC, Angelaki DE. Diverse spatial reference frames of vestibular signals in parietal cortex. *Neuron* 80: 1310–1321, 2013b.
- Chen X, DeAngelis GC, Angelaki DE. Eye-centered representation of optic flow tuning in the ventral intraparietal area. *J Neurosci* 33: 18574–18582, 2013c.
- Cohen YE, Andersen RA. A common reference frame for movement plans in the posterior parietal cortex. *Nat Rev Neurosci* 3: 553–562, 2002.
- Colby CL, Duhamel JR, Goldberg ME. Ventral intraparietal area of the macaque: anatomic location and visual response properties. *J Neurophysiol* 69: 902–914, 1993.
- Duhamel JR, Bremmer F, Ben Hamed S, Graf W. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389: 845–848, 1997.
- Duhamel JR, Colby CL, Goldberg ME. Ventral intraparietal area of the macaque: congruent visual and somatic response properties. *J Neurophysiol* 79: 126–136, 1998.
- Fetsch CR, Wang S, Gu Y, DeAngelis GC, Angelaki DE. Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. *J Neurosci* 27: 700–712, 2007.
- Gu Y, Watkins PV, Angelaki DE, DeAngelis GC. Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. *J Neurosci* 26: 73–85, 2006.
- Guipponi O, Wardak C, Ibarrola D, Comte JC, Sappey-Marinié D, Pineda S, Ben Hamed S. Multimodal convergence within the intraparietal sulcus of the macaque monkey. *J Neurosci* 33: 4128–4139, 2013.
- Lee B, Pesaran B, Andersen RA. Area MSTd neurons encode visual stimuli in eye coordinates during fixation and pursuit. *J Neurophysiol* 105: 60–68, 2011.
- Lee J, Groh JM. Auditory signals evolve from hybrid- to eye-centered coordinates in the primate superior colliculus. *J Neurophysiol* 108: 227–242, 2012.
- Lewis JW, Van Essen DC. Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. *J Comp Neurol* 428: 112–137, 2000.
- Maciokas JB, Britten KH. Extrastriate area MST and parietal area VIP similarly represent forward headings. *J Neurophysiol* 104: 239–247, 2010.
- Maier JX, Groh JM. Multisensory guidance of orienting behavior. *Hear Res* 258: 106–112, 2009.
- Mullette-Gillman OA, Cohen YE, Groh JM. Eye-centered, head-centered, and complex coding of visual and auditory targets in the intraparietal sulcus. *J Neurophysiol* 94: 2331–2352, 2005.
- Mullette-Gillman OA, Cohen YE, Groh JM. Motor-related signals in the intraparietal cortex encode locations in a hybrid, rather than eye-centered reference frame. *Cereb Cortex* 19: 1761–1775, 2009.
- Pouget A, Deneve S, Duhamel JR. A computational perspective on the neural basis of multisensory spatial representations. *Nat Rev Neurosci* 3: 741–747, 2002.
- Schlack A, Hoffmann KP, Bremmer F. Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). *Eur J Neurosci* 16: 1877–1886, 2002.
- Schlack A, Sterbing-D'Angelo SJ, Hartung K, Hoffmann KP, Bremmer F. Multisensory space representations in the macaque ventral intraparietal area. *J Neurosci* 25: 4616–4625, 2005.
- Sereno MI, Huang RS. A human parietal face area contains aligned head-centered visual and tactile maps. *Nat Neurosci* 9: 1337–1343, 2006.
- Smith MA, Majaj NJ, Movshon JA. Dynamics of motion signaling by neurons in macaque area MT. *Nat Neurosci* 8: 220–228, 2005.
- Stricanne B, Andersen RA, Mazzoni P. Eye-centered, head-centered, and intermediate coding of remembered sound locations in area LIP. *J Neurophysiol* 76: 2071–2076, 1996.
- Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE. Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. *J Neurosci* 27: 9742–9756, 2007.
- Yang Y, Liu S, Chowdhury SA, DeAngelis GC, Angelaki DE. Binocular disparity tuning and visual-vestibular congruency of multisensory neurons in macaque parietal cortex. *J Neurosci* 31: 17905–17916, 2011.
- Zhang T, Britten KH. The responses of VIP neurons are sufficiently sensitive to support heading judgments. *J Neurophysiol* 103: 1865–1873, 2010.