2011

Representation of vestibular and visual cues to self-motion in ventral intraparietal cortex

Aihua Chen
Washington University School of Medicine in St. Louis

Gregory C. DeAngelis
University of Rochester

Dora E. Angelaki
Washington University School of Medicine in St. Louis

Follow this and additional works at: http://digitalcommons.wustl.edu/open_access_pubs
Part of the Medicine and Health Sciences Commons

Behavioral/Systems/Cognitive

Representation of Vestibular and Visual Cues to Self-Motion in Ventral Intraparietal Cortex

Aihua Chen,[1] **Gregory C. DeAngelis,**[2] **and Dora E. Angelaki**[1]

[1]Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110, and [2]Department of Brain and Cognitive Sciences, Center for Visual Science, University of Rochester, Rochester, New York 14627

Convergence of vestibular and visual motion information is important for self-motion perception. One cortical area that combines vestibular and optic flow signals is the ventral intraparietal area (VIP). We characterized unisensory and multisensory responses of macaque VIP neurons to translations and rotations in three dimensions. Approximately one-half of VIP cells show significant directional selectivity in response to optic flow, one-half show tuning to vestibular stimuli, and one-third show multisensory responses. Visual and vestibular direction preferences of multisensory VIP neurons could be congruent or opposite. When visual and vestibular stimuli were combined, VIP responses could be dominated by either input, unlike the medial superior temporal area (MSTd) where optic flow tuning typically dominates or the visual posterior sylvian area (VPS) where vestibular tuning dominates. Optic flow selectivity in VIP was weaker than in MSTd but stronger than in VPS. In contrast, vestibular tuning for translation was strongest in VPS, intermediate in VIP, and weakest in MSTd. To characterize response dynamics, direction–time data were fit with a spatiotemporal model in which temporal responses were modeled as weighted sums of velocity, acceleration, and position components. Vestibular responses in VIP reflected balanced contributions of velocity and acceleration, whereas visual responses were dominated by velocity. Timing of vestibular responses in VIP was significantly faster than in MSTd, whereas timing of optic flow responses did not differ significantly among areas. These findings suggest that VIP may be proximal to MSTd in terms of vestibular processing but hierarchically similar to MSTd in terms of optic flow processing.

Introduction

Multiple cortical areas contain neurons that show selective responses to both visual motion stimuli encountered during self-motion (optic flow) (Warren, 2004; Britten, 2008) and inertial motion in darkness (which activates vestibular receptors) (Angelaki and Cullen, 2008), including the ventral intraparietal area (VIP) (Colby et al., 1993; Bremmer et al., 1999, 2002b; Zhang et al., 2004; Maciokas and Britten, 2010; Zhang and Britten, 2010), the dorsal medial superior temporal area (MSTd) (Bremmer et al., 1997; Duffy, 1998; Page and Duffy, 2003; Gu et al., 2006; Takahashi et al., 2007; Maciokas and Britten, 2010), and the visual posterior sylvian area (VPS) (Chen et al., 2011b). There is growing evidence that MSTd is involved in multisensory visual/vestibular heading perception (Britten and van Wezel, 1998; Gu et al., 2007, 2008, 2010). It remains uncertain whether this is also true for VIP.

There are multiple reasons to consider VIP an important candidate area for self-motion perception. First, like MSTd, VIP is a midlevel area in the dorsal visual processing stream (Felleman and Van Essen, 1991). Second, areas VIP and MSTd are anatomically linked (Boussaoud et al., 1990; Baizer et al., 1991). Third, many VIP neurons respond to optic flow in similar fashion to MSTd neurons (Colby et al., 1993; Schaafsma and Duysens, 1996; Duhamel et al., 1998; Bremmer et al., 2002a; Zhang and Britten, 2004, 2010; Zhang et al., 2004). Fourth, many VIP neurons are multisensory, showing convergence of visual and tactile (Colby et al., 1993; Duhamel et al., 1998; Avillac et al., 2007), visual and auditory (Schlack et al., 2005), or visual and vestibular (Bremmer et al., 2002b; Schlack et al., 2002) signals.

Area VIP receives a major visual projection from the middle temporal area (MT) (Maunsell and van Essen, 1983; Ungerleider and Desimone, 1986), tactile inputs from primary somatosensory cortex (Seltzer and Pandya, 1986), vestibular inputs from the parieto-insular vestibular cortex (PIVC) (Lewis and Van Essen, 2000), and is reciprocally connected to portions of premotor cortex responsible for head movements, oral prehension, and coordinated hand–mouth actions (Lewis and Van Essen, 2000). VIP neurons respond both to yaw rotation (Bremmer et al., 2002b; Klam and Graf, 2003) and forward–backward translation of the body (Schlack et al., 2002). However, relatively little is known about the vestibular response properties of VIP neurons and how they interact with optic flow selectivity.

We measured the three-dimensional (3D) tuning of VIP neurons to translations and rotations using experimental protocols identical with those used previously to study areas MSTd (Gu et al., 2006; Takahashi et al., 2007) and VPS (Chen et al., 2011b).

Preliminary data have been presented in abstract form (Chen et al., 2007).

Materials and Methods

Subjects and setup. Extracellular recordings were obtained from seven hemispheres in five male rhesus monkeys (*Macaca mulatta*) weighing between 6 and 10 kg. The surgical preparation, experimental apparatus, and methods of data acquisition have been described in detail previously (Gu et al., 2006; Fetsch et al., 2007; Takahashi et al., 2007). Briefly, each animal was chronically implanted with a circular molded, lightweight plastic ring for head restraint and a scleral coil for monitoring eye movements inside a magnetic field (CNC Engineering). Behavioral training was accomplished using standard operant conditioning. All surgical and experimental procedures were approved by the Institutional Animal Care and Use Committee at Washington University and were in accordance with NIH guidelines.

During experiments, the monkey was seated comfortably in a primate chair, which was secured to a 6 df motion platform (MOOG 6DOF2000E). Three-dimensional movements along or around any arbitrary axis were delivered by this platform. In all experiments, the head was positioned such that the horizontal stereotaxic plane was earth-horizontal, with the axis of rotation always passing through the center of the head (i.e., the midline point along the interaural axis). Computer-generated visual stimuli were rear-projected (Christie Digital Mirage 2000) onto a tangent screen placed ~30 cm in front of the monkey (subtending 90 × 90° of visual angle) and simulated self-motion through a three-dimensional cloud of random dots (100 cm wide, 100 cm tall, and 40 cm deep). Visual stimuli were programmed using the OpenGL graphics library and were generated by an OpenGL accelerator board (Quadro FX 3000G; PNY Technologies) (for details, see Gu et al., 2006). Monkeys viewed these visual stimuli stereoscopically as red/green anaglyphs through Kodak Wratten filters (red no. 29, green no. 61). The projector, screen, and magnetic field coil frame were mounted on the platform and moved together with the animal. Details regarding the experimental setup can be found in the studies by Gu et al. (2006, 2008) and Takahashi et al. (2007).

Tungsten microelectrodes (FHC; tip diameter, 3 μm; impedance, 1–2 MΩ at 1 kHz) were inserted into the cortex through a transdural guide tube, using a hydraulic microdrive. Behavioral control and data acquisition were accomplished by custom scripts written for use with the TEMPO system (Reflective Computing). Neural voltage signals were amplified, filtered (400–5000 Hz), discriminated (Bak Electronics), and displayed on an oscilloscope. The times of occurrence of action potentials and all behavioral events were recorded with 1 ms resolution. Raw neural signals were also digitized at a rate of 25 kHz using the CED Power 1401 system (Cambridge Electronic Design) for off-line spike sorting.

Anatomical localization. The relevant areas in the intraparietal sulcus were first identified using MRI scans. An initial ("baseline") scan was performed on each monkey before any surgeries using a high-resolution sagittal MPRAGE sequence (0.75 × 0.75 × 0.75 mm voxels). SUREFIT software (Van Essen et al., 2001) was used to segment gray matter from white matter. A second scan was performed after the head holder and recording grid had been surgically implanted. Small cannulae filled with a contrast agent (gadoversetamide) were inserted into the recording grid during the second scan to register electrode penetrations with the MRI volume. The MRI data were converted to a flat map using CARET software (Van Essen et al., 2001) and the flat map was morphed to match a standard macaque atlas (Van Essen et al., 2001). The data were then refolded and transferred onto the original MRI volume.

With the MRI scans and functional boundaries as a guide, we performed electrode penetrations to map an extensive region in and around the intraparietal cortex. In two animals (both hemispheres of monkey J and right hemisphere of monkey C), electrode penetrations were directed to the general area of gray matter around the medial tip of the intraparietal sulcus with the goal of characterizing the entire anterior–posterior extent of area VIP (typically defined as the intraparietal area with directionally selective visual responses) (Colby et al., 1993; Duhamel et al., 1997). At each location along the anterior–posterior axis, we first identified the medial tip of the intraparietal sulcus and then moved lat-

erally until there was no longer directionally selective visual response in the multiunit activity. At the anterior end, visually responsive neurons gave way to purely somatosensory cells in the fundus. At the posterior end, direction-selective neurons gave way to visual cells that were not selective for motion.

For some neurons, we also mapped receptive fields (RFs) by moving a patch of drifting random dots around the visual field and observing a qualitative map of instantaneous firing rates on a custom graphical interface. For most VIP neurons, RFs were centered in the contralateral visual field but also extended into the ipsilateral field and included the fovea. Many of the RFs were well contained within the boundaries of our display screen, but some RFs clearly extended beyond the boundaries of the screen. Moreover, VIP neurons usually were activated only by large visual stimuli (random-dot patches >10 × 10°), with smaller patches typically evoking little response.

Within the region of gray matter with robust visual motion responses, we recorded from any neuron that responded to a large-field flickering random-dot stimulus or was spontaneously active (even if it did not respond to the random-dot stimulus). Thus, there was no preselection of cells based on particular response properties. The locations of recorded neurons were reconstructed and plotted on coronal MRI sections. In two animals (monkey J, 194 cells; monkey C, 154 cells), our recordings in VIP extended over a 10 mm range anterior to posterior and included the upper bank, lower bank and tip of the intraparietal sulcus, as illustrated in Figure 1. For the other three animals, electrode penetrations were restricted to smaller regions in VIP (animal U, 14 cells; animal O, 54 cells; animal P, 36 cells). To allow a direct comparison between the response properties of VIP and MSTd/VPS neurons, we have also included data from 336 MSTd cells (animal A, 22 cells; animal Q, 124 cells; animal Z, 190 cells) and 166 VPS cells (animal E, 116 cells; animal A, 50 cells) tested with vestibular and visual translation stimuli, as well as 128 MSTd cells tested with rotation stimuli (animal A, 27 cells; animal L, 37 cells; animal Q, 59 cells; animal Z, 5 cells) (for details of the MSTd and VPS studies, see Gu et al., 2006, 2010; Fetsch et al., 2007; Takahashi et al., 2007; Chen et al., 2011b). Note that the sample of VIP neurons was recorded from different animals than the samples of MSTd and VPS neurons.

Experimental protocol. Once action potentials from a single VIP neuron were satisfactorily isolated, regardless of the strength of its visual or vestibular activity, responses were first characterized during a 3D translation protocol (Gu et al., 2006). Stimuli were presented along 26 heading directions sampled from a sphere, corresponding to all combinations of azimuth and elevation angles in increments of 45° (see Fig. 2A). This included all combinations of movement vectors having 8 different azimuth angles (0, 45, 90, 135, 180, 225, 270, and 315°, where 0 and 90° correspond to rightward and forward translations, respectively) and 3 different elevation angles: 0° (the horizontal plane) and ±45° (for a total of 8 × 3 = 24 directions). In addition, elevation angles of −90 and 90° were included to generate upward and downward movement directions, respectively. The duration of the motion stimulus was 2 s, its velocity profile was Gaussian, and it had a corresponding biphasic acceleration profile. The motion amplitude was 13 cm (total displacement), with a peak acceleration of ~0.1 g (~0.98 m/s^2) and a peak velocity of ~30 cm/s (see Fig. 2B).

Within a single block of trials, either two or three stimulus types were randomly interleaved: (1) In the "vestibular" condition, the monkey was physically translated by the motion platform along each of the 26 possible heading trajectories, in the absence of optic flow. The screen was blank, except for a head-centered fixation point. Note that we refer to this stimulus as the vestibular condition for simplicity, although other nonvisual contributions (e.g., somatosensory or proprioceptive) cannot be excluded. (2) In the "visual" condition, the motion platform was stationary while optic flow presented on the display screen simulated movement through a cloud of stars along the same set of 26 possible headings. Note that all stimulus directions are referenced to body motion (real or simulated). Thus, a neuron with similar direction preferences in the visual and vestibular conditions would be considered "congruent." (3) In the "combined" condition, the animal was translated by the motion platform while a congruent optic flow stimulus was simultaneously presented. Visual and vestibular stimuli were precisely synchronized and visual

stimuli contained binocular disparity, motion parallax, and size cues that specified the scale of the visual environment (for details, see Gu et al., 2006).

Once the 3D translation protocol was completed and if good cell isolation was maintained, most VIP neurons recorded from monkey J and a few cells recorded from monkey U were then tested with a 3D rotation protocol. Other VIP neurons were instead tested using a heading discrimination task (monkey U and monkey C) or a disparity task (monkey O and monkey P), as part of other ongoing studies. Thus, we only attempted to conduct the rotation protocol on a subset (216 of 452) of the neurons for which we had completed the translation protocol. In the rotation protocol, stimulus directions were defined by the same set of 26 vectors, which now represent the corresponding axes of rotation according to the right hand rule (Takahashi et al., 2007). For example, azimuth angles of 0 and 180° (elevation, 0°) correspond to pitch-up and pitch-down rotations, respectively. Azimuths of 90 and 270° (elevation, 0°) correspond to roll rotations (right-ear-down and left-ear-down, respectively). Finally, elevation angles of −90 or 90° correspond to leftward and rightward yaw rotation, respectively. The rotational motion trajectory followed a Gaussian velocity profile and rotation amplitude was 9° (peak angular velocity, ~20°/s).

During all stimulus conditions (visual, vestibular, or combined; translation or rotation), the animal was required to fixate a central target (0.2° in diameter) for 200 ms before stimulus onset (fixation windows spanned 2 × 2° of visual angle). The fixation target moved with the head (head-centered) such that no eye movement was required to maintain gaze on the target. The animals were rewarded at the end of each trial for maintaining fixation throughout stimulus presentation. If fixation was broken at any time during the stimulus, the trial was aborted and the data were discarded. Neurons were included in the sample if each stimulus was successfully repeated at least three times. Across our sample of VIP neurons, 90% of cells were isolated long enough for at least five stimulus repetitions.

Within a block of trials, visual and vestibular stimuli were randomly interleaved along with a (null) condition in which the motion platform remained stationary and no star field was shown (to assess spontaneous activity). To complete five repetitions of all 26 directions for the visual and vestibular conditions, plus five repetitions of the null condition, the monkey was required to successfully complete 26 × 2 × 5 + 5 = 265 trials. For a subpopulation of cells (translation, n = 50; rotation, n = 30), combined visual/vestibular stimuli were interleaved, for a total of 395 trials. Translation and rotation stimuli were presented in separate blocks of trials. Note that these protocols are identical with those used previously to characterize MSTd (Gu et al., 2006; Takahashi et al., 2007) and PIVC neurons (Chen et al., 2010).

For a subpopulation of VIP neurons that showed significant tuning in the vestibular condition, neural responses were also collected during platform motion along each of the 26 different directions in complete darkness (with the projector turned off). In these controls, there was no behavioral requirement to fixate and rewards were delivered manually to keep the animal motivated and alert.

Data analysis. Analysis of spike data and statistical tests were performed using MATLAB (MathWorks). Peristimulus time histograms (PSTHs) were constructed for each direction of translation/rotation using 25 ms time bins and were smoothed with a 400 ms sliding boxcar filter. The temporal modulation of the response for each stimulus direction was considered significant when the spike count distribution from the time bin containing the maximum and/or minimum response differed significantly from the baseline response distribution (−100 to 300 ms after stimulus onset; Wilcoxon's matched-pairs test, $p < 0.01$) (for details, see Chen et al., 2010). We calculated the maximum response of the neuron across stimulus directions for each 25 ms time bin between 0.5 and 2 s after motion onset. We then used ANOVA to assess the statistical significance of direction tuning as a function of time and to evaluate whether there are multiple time periods in which a neuron shows distinct temporal peaks of directional tuning (for details, see Chen et al., 2010). "Peak times" were defined as the times of local maxima at which distinct epochs of directional tuning were observed, using, the following criteria: (1) ANOVA values are significant ($p < 0.01$) for five

consecutive time bins centered on the putative local maximum, and (2) there is a continuous temporal sequence of time bins for which direction tuning is significantly positively correlated with tuning at the peak time.

Based on the number of distinct peak times, VIP cells were divided into three groups: (1) cells with a single temporal epoch of directional selectivity ("single-peaked"), (2) cells with two distinct epochs of direction tuning with a gap in between ("double-peaked"), and (3) cells that were not significantly direction-selective in any time period ("not-tuned"). To visually represent 3D tuning for translation or rotation, mean responses are plotted as a function of azimuth and elevation. For this purpose, data are transformed using the Lambert cylindrical equal-area projection and plotted on Cartesian axes (Gu et al., 2006). This produces flattened representations in which the abscissa represents azimuth angle, and the ordinate corresponds to a sinusoidally transformed version of elevation angle. Note that the color scale in each contour plot was chosen based on the minimum and maximum responses of each neuron, rounded to the nearest multiple of 10 spikes/s.

The strength of directional tuning at each peak time was quantified using a direction discrimination index (DDI; Takahashi et al., 2007) given by:

$$DDI = \frac{R_{\max} - R_{\min}}{R_{\max} - R_{\min} + 2\sqrt{SSE/(N - M)}}, \tag{1}$$

where $R_{\max}$ and $R_{\min}$ are the maximum and minimum responses from the 3D tuning function, respectively. SSE is the sum squared error around the mean response, N is the total number of observations (trials), and M is the number of stimulus directions ($M = 26$). The DDI compares the difference in firing between the preferred and null directions against response variability, and quantifies the reliability of a neuron for distinguishing between preferred and null motion directions. Neurons with large response modulations relative to the noise level will have DDI values close to 1, whereas neurons with weak response modulation will have DDI values close to 0. DDI is conceptually similar to a d' metric in that it quantifies signal-to-noise ratio, but it has the advantage of being bounded between 0 and 1, similar to other conventional metrics of response modulation. A particular value of DDI does not correspond to a specific ratio of preferred/null direction responses because DDI also depends on the variability of responses.

The preferred direction of a neuron for each stimulus condition was described by the azimuth and elevation of the vector sum of the individual responses (after subtracting spontaneous activity). In such a representation, the mean firing rate in each trial was considered to represent the magnitude of a 3D vector whose direction was defined by the azimuth and elevation angles of the particular stimulus. Preferred directions have been plotted on Cartesian axes using the Lambert projection, as described above. This transformation was also used to calculate the distribution of the absolute difference in 3D direction preferences (|Δ preferred direction|), such that random combinations of direction preferences across stimulus conditions would yield a flat distribution.

Note that the vector sum can reliably reflect the tuning preference of the cell only when the directional tuning profile is unimodal. Thus, all analyses using the vector sum computation of direction preference were applied only to cells with unimodal tuning. Unimodality was assessed by interpolating the 3D tuning data to 5° resolution in both azimuth and elevation and applying a multimodality test based on the kernel density estimate method (Silverman, 1981; Anzai et al., 2007; Chen et al., 2010). The test generated a probability value (p_{uni}): if $p_{uni} > 0.05$, the tuning was classified as unimodal; if $p_{uni} < 0.05$, unimodality was rejected, and the tuning was classified as multimodal and not included in any vector sum calculations.

For other properties presented as distributions, a resampling analysis was used to assess whether the distribution was significantly different from uniform, as follows (Takahashi et al., 2007). We computed the sum squared error (across bins) between the measured distribution and an ideal uniform distribution containing the same number of observations. Then we generated a random distribution by drawing the same number of data points from a uniform distribution using the "unifrnd" function in MATLAB. The sum squared error was again calculated between this

random distribution and the ideal uniform distribution. This second step was repeated 1000 times to generate a distribution of sum squared error values that represent random deviations from an ideal uniform distribution. If the sum squared error for the experimentally measured distribution lay outside the 95% confidence interval of values from the randomized distributions, then the measured distribution was considered to be significantly different from uniform ($p < 0.05$).

For nonuniform distributions, the number of modes was further assessed using a multimodality test based on the kernel density estimate method (for details, see Takahashi et al., 2007). A von Mises function (the circular analog of the normal distribution) was used as the kernel for circular data. Watson's U^2 statistic, corrected for grouping, was computed as a goodness-of-fit test statistic to obtain a p value through a bootstrapping procedure. This test generates two p values, with the first one (p_{uni}) for the test of unimodality and the second one (p_{bi}) for the test of bimodality.

To quantify the visual and vestibular contributions to the combined response, we measured a "vestibular gain" and a "visual gain" by fitting the combined responses with a weighted sum of the visual and vestibular responses as follows:

$$R_{combined} = a_1 \times R_{vestibular} + a_2 \times R_{visual} + a_3. \quad (2)$$

In this equation, $R_{combined}$ is the recorded VIP response during the combined condition, which is represented as a matrix of mean firing rates for all heading directions (after subtraction of spontaneous activity). $R_{vestibular}$ and R_{visual} represent responses in the vestibular and visual conditions, respectively. Coefficients a_1 and a_2 are the vestibular and visual gains, respectively; and a_3 is a constant that accounts for direction-independent differences between the three conditions. A "gain ratio" was defined as a_1/a_2: the higher the gain ratio, the greater the vestibular contribution (relative to visual) to the combined response (Takahashi et al., 2007).

To test whether the gain ratio correlated with the relative strength of tuning in the single-cue conditions, we also computed a visual-vestibular response ratio (VVR) defined as follows (Takahashi et al., 2007):

$$VVR = \frac{R_{max}(vis) - R_{min}(vis)}{R_{max}(ves) - R_{min}(ves)}, \quad (3)$$

where $R_{max}(vis)$ and $R_{max}(ves)$ are the maximum firing rates in the visual and vestibular conditions, and $R_{min}(vis)$ and $R_{min}(ves)$ are the minimum firing rates, respectively.

Spatiotemporal curve fitting. To provide a compact description of the spatiotemporal tuning properties of neurons in areas VIP, MSTd, and VPS, we used a model to fit the heading tuning functions of these neurons (for details, see Chen et al., 2011a). To reduce the dimensionality of the fits, the model was fit to a subset of conditions corresponding to one of three cross-sections through the spherical data (eight stimulus directions in each cross-section): the horizontal plane, the vertical (frontoparallel) plane, or the median plane. Fitting was performed if there was significant space–time structure in a particular plane ($p < 0.001$, two-way ANOVA, significant main effects of space and time and a significant interaction).

Three different models were used to fit the spatiotemporal response profiles of each cell. The simplest "velocity" model (model V) consisted of the product of a modified cosine function of stimulus direction and a Gaussian velocity profile in time, as given by the following:

$$R(\theta,t) = A \times ([F(\cos(\theta - \theta_0),n)] - DC) \times G(t) + R_0. \quad (4)$$

Here, $R(\theta, t)$ represents the response of the neuron (in spikes/second) as a function of direction and time, A is the overall response amplitude, θ denotes stimulus direction (range, 0 to 2π), θ_0 represents the direction preference of the cell, DC indicates a baseline shift of the spatial tuning (range, $0–0.5$), $G(t)$ represents the temporal response profile of the neuron as defined below, and R_0 is a constant corresponding to the resting firing rate of the neuron.

The cosine tuning for direction of motion in the model is modified by two nonlinearities. First, $F(x,n)$ indicates a "squashing" nonlinearity as given by Nguyenkim and DeAngelis (2003):

$$F(x,n) = \frac{\exp(nx) - 1}{n}, n > 0. \quad (5)$$

When n approaches zero, $F(x,n) \sim x$ and the nonlinearity has no effect. As n increases, $F(x,n)$ causes the peak of the cosine function to become taller and narrower while the trough becomes shallower and wider. This nonlinearity was useful to fit the direction tuning of many neurons, which is typically somewhat narrower than pure cosine tuning (Gu et al., 2010). Second, the operation denoted by [ ] indicates that the spatial tuning curve was normalized to be in the range from [0 1] following application of $F(x,n)$. This normalization reduces correlations between the parameter of the squashing nonlinearity (n) and the overall amplitude (A) and baseline response (R_0) parameters, thus improving the convergence of the fits.

The temporal response profile in model V, $G(t)$, is a temporal Gaussian function, given by the following:

$$G(t) = e^{\left(\frac{-(t - t_0)^2}{2\sigma_t^2}\right)}, \quad (6)$$

where t_0 is the time at which the peak response occurs and σ_t indicates the SD. In total, model V has seven free parameters: A, DC, θ_0, n, σ_t, t_0, and R_0.

For each cardinal plane cross-section through the spherical data, Equation 4 was fit to the data using a nonlinear least-squares optimization procedure ("lsqcurvefit" function in MATLAB). The data to be fit were response PSTHs, having 100 ms bins, for each of the eight stimulus directions (45° apart) in one of the cardinal planes.

The "velocity plus acceleration" model (model VA) incorporates a second spatiotemporal response component to account for the possibility that both velocity and acceleration were encoded by the neurons. This additional component is the product of an offset spatial tuning curve and an acceleration profile in time [derivative of $G(t)$]. Thus, model VA allows the temporal responses to be mixtures of velocity and acceleration, and also allows the directional tuning of the acceleration component to differ from that of the velocity component. Model VA is formulated as follows:

$$R = A \times \left(\begin{array}{c} w_v \times ([F(\cos(\theta - \theta_0),n)] - DC) \times G(t) \\ + (1 - w_v) \times ([F(\cos(\theta - \theta_0 - \Delta\theta_{va}),n)] \\ - DC) \times \left[\frac{dG(t)}{dt}\right]_{-1} \end{array} \right) + R_0.$$

$$(7)$$

Model VA contains two additional free parameters compared with model V, for a total of nine. The first additional parameter is the difference between direction preferences for the velocity and acceleration components, denoted by $\Delta\theta_{va}$. The second additional parameter, w_v, specifies the relative weight of the velocity component, with the acceleration weight given by $w_a = (1 - w_v)$. For a purely velocity-driven response, $w_v = 1$; for a purely acceleration-driven response, $w_v = 0$; and for a balanced mixture, $w_v = 0.5$. A ratio of acceleration to velocity components was computed as $w_a/w_v = (1 - w_v)/w_v$. The temporal derivative term in Equation 7 was normalized into the range $[-1\ 1]$, as denoted by $[]_{-1}$.

Finally, the "velocity plus acceleration plus position" model (model VAP) incorporated an additional spatiotemporal component to allow for a position contribution. The position component was represented as the product of an offset spatial tuning curve and a position profile in time [integral of $G(t)$], as follows:

$$R = A \times \left(\begin{array}{c} w_v \times ([F(\cos(\theta - \theta_0),n)] - DC) \times G(t) \\ + (1 - w_v) \times ([F(\cos(\theta - \theta_0 - \Delta\theta_{va}),n)] \\ - DC) \times \left[\frac{dG(t)}{dt}\right]_{-1} \\ + w_p \times ([F(\cos(\theta - \theta_0 - \Delta\theta_{vp}),n)] - DC) \times [\int G(t)dt] \end{array} \right)$$

$$+ R_0. \quad (8)$$

Relative to model VA, model VAP again adds two additional free parameters (for a total of 11). One parameter is the difference in direction preference between the velocity and position components, $\Delta\theta_{vp}$, and the second parameter is the weight of the position component, w_p, ranging from 0 to 1. Note that the temporal integral term in Equation 8 was normalized into the range [0 1], as done for the velocity component.

In all of these models, $t_0 = 1$ s corresponds to the peak of the Gaussian velocity profile of the stimulus. In the model fits presented, all parameters were free to take any value, except for DC, which was bounded in the range [0 0.5].

As the number of free parameters differed between the three models, model selection was performed using a sequential F test. A significant outcome of the sequential F test ($p < 0.01$) indicates that the model with a greater number of parameters outperforms the simpler model (allowing for the extra degrees of freedom). We chose the sequential F test, rather than the Akaike information criterion used previously (Chen et al., 2011a), because we found the F test to be more conservative in terms of accepting greater model complexity, and the results of the sequential F test were more acceptable by eye, especially for visual responses.

Results

Quantitative data were obtained from 452 VIP neurons recorded from seven hemispheres in five rhesus monkeys. The majority of neurons were recorded from the right and left hemispheres of animal J and the right hemisphere of animal C, as illustrated in Figure 1, black symbols (cells with significant responses to visual and/or vestibular translation) and white symbols (cells without significant responses to translation). Cells responsive to translation were encountered in both the upper and lower banks of the intraparietal sulcus, with recordings concentrated around the medial tip of the sulcus. Upon isolation, each VIP neuron was first tested with physical (vestibular condition) and simulated (visual condition) translation along 26 motion directions uniformly distributed in 3D space (Fig. 2A). If satisfactory isolation was maintained throughout this translation protocol, some cells (see Materials and Methods) were subsequently tested with physical and simulated rotation about the same 26 axes, with each axis defining a direction of rotation according to the right-hand rule. Each movement followed a Gaussian velocity profile, with a corresponding biphasic acceleration profile (Fig. 2B). For each block of trials (translation or rotation), visual and vestibular conditions were randomly interleaved, along with a null condition in which monkeys fixated the head-centered target without any visual or vestibular stimulation. We begin by describing the properties of VIP responses to translation.

Visual and vestibular responses to translation

Typical responses from a "congruent" cell and an "opposite" cell are illustrated in Figure 2, C and D, respectively. Both of these neurons have single-peaked tuning, with a single significant epoch of directional selectivity. The plots on the left show average
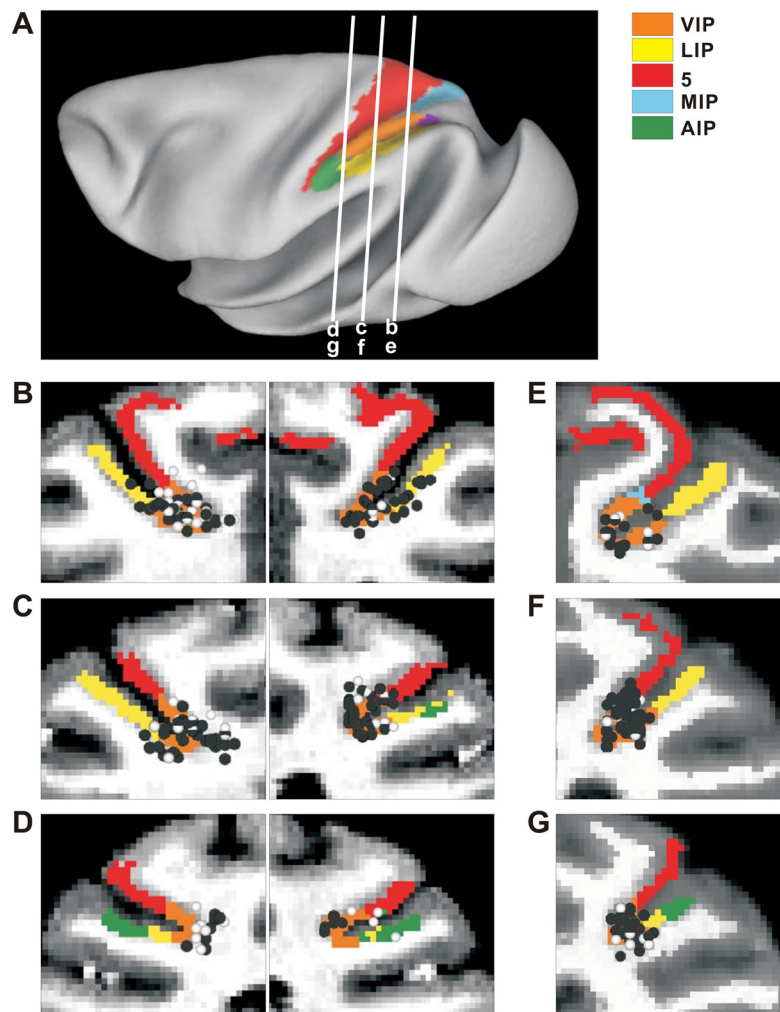
Figure 1. Anatomical localization of recording sites. **A**, Inflated cortical surface illustrating the locations of the coronal sections drawn in **B–G**. **B–D**, Coronal sections from both hemispheres of monkey J, spaced 4 mm apart, are shown from posterior (**B**) to anterior (**D**). **E–G**, Coronal sections from the right hemisphere of monkey C are shown from posterior (**E**) to anterior (**G**). Cells located within 2 mm of each section were projected onto that section. The black symbols represent single units with significant tuning to either vestibular or visual translation. The white symbols represent cells that showed no directional tuning.

PSTHs for all 26 directions of vestibular (top) and visual (bottom) translation, arranged according to stimulus direction in spherical coordinates. The red dashed lines show the peak response times for each neuron under each stimulus condition. A peak time is defined as the center of the 400 ms time window that produces the largest departure in firing rate from the baseline response (see Materials and Methods).

The 3D directional tuning of these example neurons, computed at the corresponding peak times, is shown as color contour maps (elevation vs azimuth) on the right side of Figure 2. The preferred direction of each neuron was defined as the azimuth and elevation of the vector sum of the neural responses (see Materials and Methods). The congruent cell of Figure 2C has similar direction preferences for vestibular and visual translation stimuli; [azimuth, elevation] = [−24, −38°] and [−6, −64°], respectively, corresponding to an upward and slightly rightward trajectory. In contrast, the opposite cell in Figure 2D preferred a backward–rightward direction in the vestibular condition ([azimuth, elevation] = [−64, 7°]) and a forward–leftward direction in the visual condition ([azimuth, elevation] = [129, 7°]). As described further below, congruent and opposite cells were both frequently encountered in VIP.
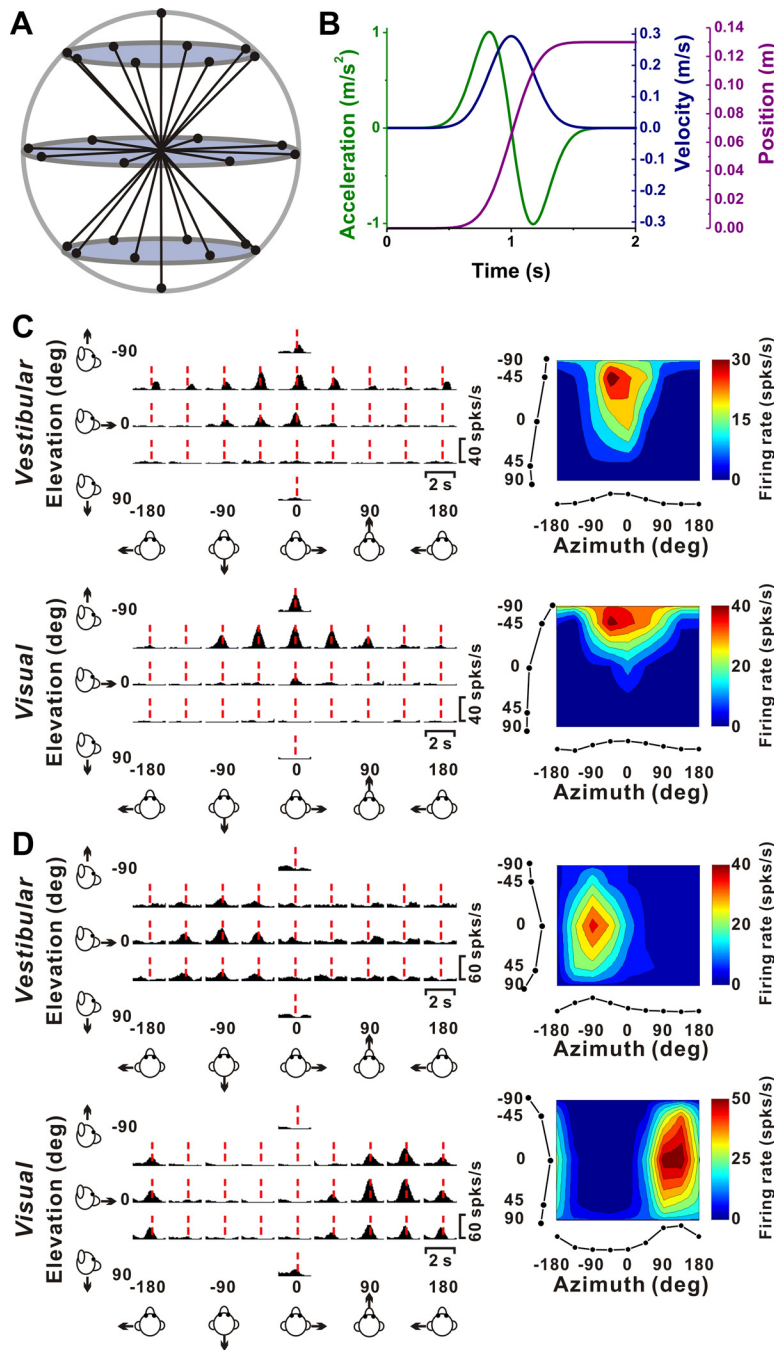
Figure 2. Stimuli and examples of 3D translation tuning. **A**, Schematic of the 26 movement trajectories in 3D, spaced 45° apart in both azimuth and elevation. **B**, The 2 s translational motion stimulus: velocity (blue), acceleration (green), and position (magenta). **C**, Response PSTHs (left panels) and 3D tuning profiles (right panels) for a congruent VIP neuron. The red line indicates the peak time ($t_{vestibular} = 1.04$; $t_{visual} = 0.91$ s) when the maximum response across directions occurred. The 3D tuning profile (right) is illustrated as a color contour map (Lambert cylindrical projection), taken at the peak response time (vestibular DDI, 0.77; visual DDI, 0.87). Tuning curves along the margins of the color map illustrate mean firing rates plotted versus elevation or azimuth (averaged across azimuth or elevation, respectively). The preferred directions for this cell (computed as vector sum) are [azimuth, elevation] = [−24, −38°] for the vestibular condition, and [−6, −64°] for the visual condition. **D**, PSTHs and spatial tuning profile for a VIP neuron with opposite direction preferences in the vestibular ([azimuth, elevation] = [−64, 7°]; DDI, 0.64; peak time, 0.94 s) and visual conditions ([129, 7°]; DDI, 0.75; peak time, 1.06 s).

Figure 3 illustrates another example VIP cell, which was classified as "double-peaked" based on its vestibular translation tuning. As illustrated by the vestibular PSTHs (Fig. 3A), there were two peak times for this cell, one at 0.91 s (red line) and another at 1.41 s (green line). The cell showed significant directional tuning at both peak times, with nearly opposite direction preferences at

[azimuth, elevation] = [39, −16°] and [−126, 9°], respectively (Fig. 3C,D). This occurs because the spatiotemporal response profile of the cell has two distinct temporal epochs of direction tuning. The early peak time occurs just before peak stimulus velocity, whereas the second peak time occurs after the second peak of the stimulus acceleration profile. Note that the visual response of this same neuron was single-peaked (Fig. 3B,E).

Among 452 cells tested in VIP, 222 (49%) showed significant tuning for vestibular translation, compared with 50–60% in MSTd (Gu et al., 2006; Takahashi et al., 2007; Liu and Angelaki, 2009). Unlike in MSTd, where >95% of the cells show significant tuning for visual translation simulated by optic flow (Gu et al., 2006), only 310 of 452 (69%) VIP cells showed significant directional tuning for visual translation. Of these, 182 cells (40%) were significantly tuned ($p < 0.01$, ANOVA) to both visual and vestibular translation, 128 cells (28%) were tuned to visual stimuli only, and 40 cells (9%) were tuned to vestibular stimuli only (such cells with only vestibular tuning are extremely rare in MSTd). All of these proportions were significantly greater than expected by chance ($p < 0.001$, χ^2 test). As shown in Table 1, only 3% of visual responses to translation in VIP were classified as double-peaked, whereas ~21% of the vestibular translation responses were double-peaked. Thus, single-peaked responses predominated.

Cells with significant directional tuning were further subdivided based on whether the spatial tuning (at a particular peak time) was unimodal or multimodal (see Materials and Methods) (Chen et al., 2010). The vast majority of VIP neurons showed unimodal direction tuning: 83% (184 of 222) for vestibular translation and 77% (240 of 310) for visual translation. Direction preferences of VIP neurons with significant unimodal tuning were distributed throughout the spherical stimulus space, as illustrated in Figure 4, A (vestibular translation) and B (visual translation). Each data point in these scatter plots specifies the preferred 3D direction of a single neuron (black, multisensory cells; red, vestibular only; green, visual only), while histograms along the boundaries show the marginal distributions of azimuth and elevation preferences. As in MSTd (Gu et al., 2006, 2010), the distribution of azimuth preferences for VIP visual responses was significantly bimodal ($p \ll 0.001$, uniformity test; $p_{uni} < 0.001$, $p_{bi} = 0.97$, modality test) (see Materials and Methods), with most cells preferring lateral over forward–backward directions. Among 240 cells with significant
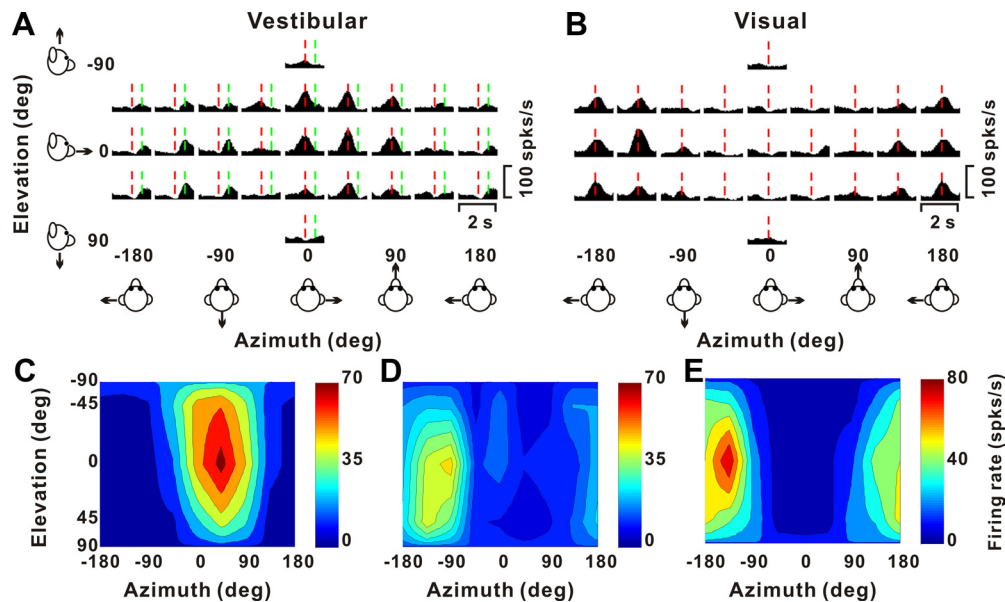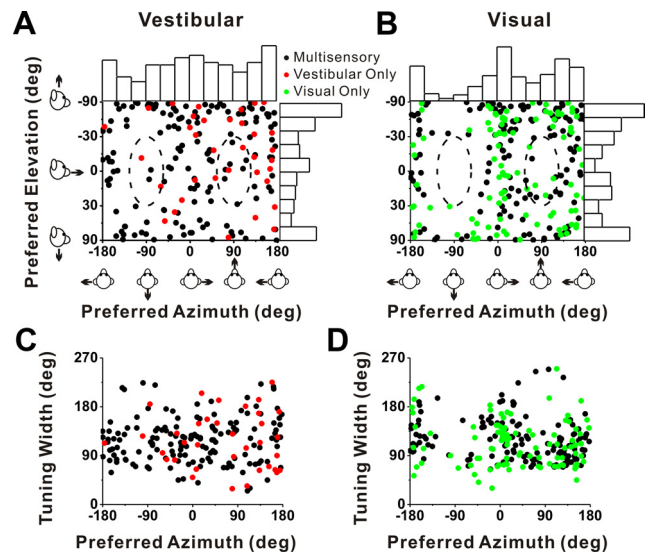
Figure 3. Example of a VIP neuron with double-peaked direction tuning in the vestibular condition. **A**, **B**, PSTHs for 26 directions during presentation of vestibular (**A**) and visual (**B**) translation stimuli. The red and green lines indicate the two peak times for the vestibular translation response ($t_{vestibular}$ = 0.91 and 1.41 s). Note that there was only one peak time for the visual response (t_{visual} = 0.96 s; red). **C, D**, The 3D directional tuning for the vestibular condition is illustrated at each of the two peak times. At the first peak time (**C**), the direction preference is [azimuth, elevation] = [39, −16°], and the DDI is 0.84; at the second peak time (**D**), the direction preference is [−126, 9°], and the vestibular DDI is 0.78. **E**, Three-dimensional directional tuning for the visual condition. The preferred direction is [−168, 12°], and the visual DDI is 0.79.

Table 1. Classification of tuned cells as single-peaked and double-peaked

	Tuned			
	Single-peaked [n (%)]	Double-peaked [n (%)]	Triple-peaked [n (%)]	Total
Translation				
VIP				
Vestibular	176 (79.3)	46 (20.7)	0 (0)	222
Visual	301 (97.1)	9 (2.9)	0 (0)	310
Combined	36 (87.8)	5 (12.2)	0 (0)	41
MSTd				
Vestibular	160 (80.0)	40 (20.0)	0 (0)	200
Visual	306 (94.7)	17 (5.3)	0 (0)	323
Combined	220 (92.8)	17 (7.2)	0 (0)	237
VPS				
Vestibular	63 (52.5)	51 (42.5)	6 (5.0)	120
Visual	56 (84.8)	10 (15.2)	0 (0)	66
Combined	23 (50.0)	21 (45.6)	2 (4.4)	46
Rotation				
VIP				
Vestibular	61 (96.8)	2 (3.2)	0 (0)	63
Visual	59 (98.3)	1 (1.7)	0 (0)	60
Combined	22 (91.7)	2 (8.3)	0 (0)	24
MSTd				
Vestibular	91 (79.8)	23 (20.2)	0 (0)	114
Visual	114 (91.9)	10 (8.1)	0 (0)	124
Combined	24 (96.0)	1 (4.0)	0 (0)	25
VPS				
Vestibular	38 (86.4)	6 (13.6)	0 (0)	44

Figure 4. Summary of direction tuning properties of VIP neurons during translation. **A**, **B**, Distribution of vestibular (**A**) and visual (**B**) 3D heading preferences. Each data point in the scatter plot corresponds to the preferred azimuth (abscissa) and elevation (ordinate) of a single neuron with significant unimodal heading tuning (**A**, n = 184; **B**, n = 240). The data are plotted on Cartesian axes that represent the Lambert cylindrical equal-area projection of the spherical stimulus space. Histograms along the top and right sides of each scatter plot show the marginal distributions. The dashed elliptical curves represent a ±30° range of directions around straight forward ([azimuth, elevation] = [90, 0°]) and straight backward ([azimuth, elevation] = [−90, 0°]). **C, D**, Scatter plots of the tuning width at half-maximum of each cell, computed from the tuning curve in the horizontal plane, versus preferred azimuth. The black dots represent cells with significant unimodal spatial tuning during both the vestibular (n = 149) and visual (n = 141) conditions. The red dots represent cells with significant unimodal vestibular tuning only (n = 35). The green dots represent cells with significant unimodal visual tuning only (n = 99).

visual translation tuning, none had direction preferences within ±30° of the backward axis and only 11 cells (4.6%) had preferred directions within ±30° of straight forward (Fig. 4B, dashed oval regions).

A similar tendency toward a bimodal distribution of azimuth preferences was seen for vestibular translation, but the effect was not significant. The distribution of azimuth preferences was not significantly different from uniform (p = 0.35, uniformity test), although slightly more cells preferred lateral and vertical move-

ments. Only 4 of 184 (2.2%) cells with significant vestibular translation tuning had a preferred direction within ±30° of straight backward and 9 cells (4.9%) had a preferred direction within ±30° of straight forward (Fig. 4A, dashed oval regions).
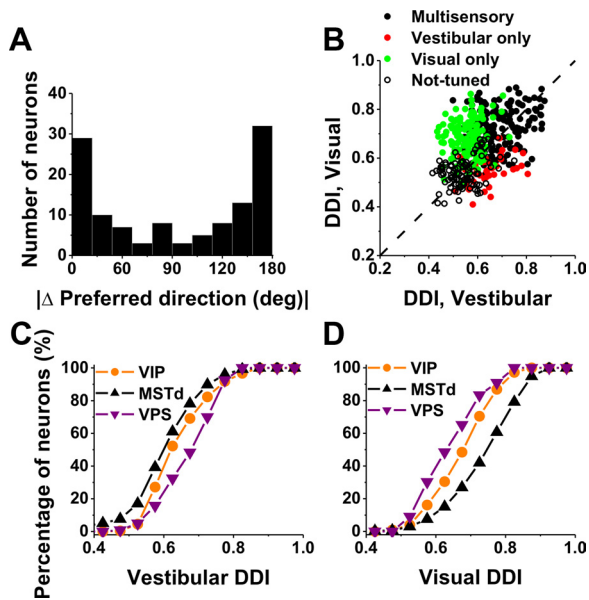
Figure 5. Comparison of tuning preferences and direction selectivity between visual and vestibular responses of VIP neurons during translation. ***A***, Distribution of the absolute difference in 3D preferred direction (|Δ preferred direction|) between visual and vestibular responses ($n = 118$). Note that bins were computed according to the cosine of the angle (in accordance with the spherical nature of the data, such that the distribution would be flat if there were no systematic relationship between visual and vestibular direction preferences). Only neurons with significant unimodal spatial tuning during both vestibular and visual conditions have been included. ***B***, Scatter plot of the visual DDI as a function of the vestibular DDI. Black filled symbols, Cells with significant tuning during both the vestibular and visual conditions ($n = 182$); red symbols, cells with significant tuning during the vestibular condition only ($n = 40$); green symbols, cells with significant tuning during the visual condition only ($n = 128$); open symbols, cells without significant tuning in either condition ($n = 102$). Dashed line, Unity-slope diagonal. ***C***, Cumulative distributions of DDI for vestibular responses to translation in VIP (orange; $n = 452$), MSTd (black; $n = 336$), and VPS (purple; $n = 166$). ***D***, Cumulative distributions of DDI for visual responses.

The distributions of elevation preference are significantly non-uniform, but unimodal, for both visual and vestibular responses ($p < 0.01$, uniformity test; $p_{uni} > 0.28$, modality test).

As illustrated in Figure 4, *C* and *D*, which shows the tuning width at half-maximum (in the horizontal plane) from a spline fit (interpolated at the 1° resolution) versus azimuth preference for each neuron, tuning was generally broad independently of azimuth preference, and tuning width was similar for unisensory and multisensory cells. Compared with area MSTd (Gu et al., 2006, 2010), we found that the average tuning width in VIP was significantly narrower for the visual translation condition ($p = 0.001$, Wilcoxon's rank sum test). However, there was no significant difference in tuning width between areas for the vestibular translation condition ($p = 0.17$).

As shown by the examples of Figure 2, some VIP cells have similar direction preferences for visual and vestibular stimuli, whereas others have opposite preferences. The distribution of the absolute difference in direction preference (|Δ preferred direction|) between visual and vestibular stimulus conditions was strongly bimodal ($p = 0.007$, uniformity test; $p_{uni} \ll 0.001$, $p_{bi} = 0.53$, modality test), as illustrated in Figure 5*A*. As found previously for translation responses in MSTd (Gu et al., 2006), congruent and opposite neurons were encountered in roughly equal proportions in VIP: 37% (44 of 118) had |Δ preferred direction| $< 60°$, and 42% (50 of 118) had |Δ preferred direction| $> 120°$. In subsequent figures, for simplicity, we refer

to cells with |Δ preferred direction| $< 90°$ as congruent and those with |Δ preferred direction| $> 90°$ as opposite.

The strength of direction tuning of VIP neurons was quantified using a DDI, which ranges from 0 (poor tuning) to 1 (strong tuning). DDI values for visual and vestibular translation conditions are compared in Figure 5*B*. Considering all cells ($n = 452$), the vestibular DDI was significantly smaller than the visual DDI (Wilcoxon's matched-pairs test, $p < 0.001$). In addition, multisensory neurons were more strongly tuned, on average, than unisensory neurons. Specifically, the vestibular DDI of multisensory neurons (0.68 ± 0.01 SE) was significantly greater than the vestibular DDI for vestibular-only neurons (0.64 ± 0.01) ($p = 0.01$, Wilcoxon's rank sum test) (Fig. 5*B*, compare black symbols, red symbols). Similarly, the visual DDI of multisensory neurons (0.72 ± 0.01 SE) was significantly greater than the visual DDI for visual-only neurons (0.68 ± 0.01) ($p < 0.001$, Wilcoxon's rank sum test) (Fig. 5*B*, compare black symbols, green symbols).

Because the experimental protocols and cell sampling criteria used here were identical with those previously used to characterize optic flow and vestibular tuning in MSTd and VPS, a direct comparison between areas is possible. These comparisons are summarized graphically (Fig. 5*C*,*D*) as cumulative distributions of DDI for VIP (orange), MSTd (black), and VPS (purple), shown separately for visual and vestibular responses, respectively. Overall, the vestibular translation DDI for VIP (mean $\pm$ SE, 0.61 ± 0.01) was modestly but significantly greater than that for MSTd (0.59 ± 0.01) ($p = 0.01$, Wilcoxon's rank sum test) but weaker than that for VPS (0.69 ± 0.01) ($p < 0.01$). In contrast, the visual translation DDI for VIP (0.66 ± 0.01) was substantially less than that for MSTd (0.76 ± 0.01) ($p < 0.001$, Wilcoxon's rank sum test) but greater than that for VPS (0.60 ± 0.01) ($p < 0.001$). Similar results were obtained when tuning strength was quantified as the vector sum of responses across directions or as the raw difference in firing rate between stimuli that elicited maximal and minimal responses (data not shown). Thus, whereas MSTd neurons are more selective to optic flow and VPS neurons are more selective to vestibular inputs, heading tuning in VIP is relatively balanced across the two modalities.
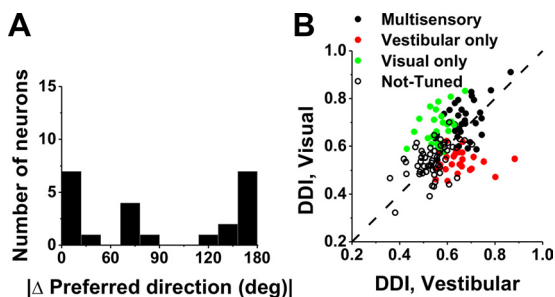
Visual and vestibular responses to rotation
We now turn to the subset of VIP cells (216 of 452) that was also tested with rotation stimuli. Note that this subset is relatively small because we did not attempt the rotation protocol on some neurons, but rather ran protocols for other studies (see Materials and Methods). Among 142 cells for which isolation was maintained and the rotation protocol was completed, 63 (44%) were significantly tuned ($p < 0.01$, ANOVA) to vestibular rotation and 60 (42%) were significantly tuned to visual rotation. Approximately one-quarter (36 cells) were multisensory neurons, whereas 24 cells (17%) were tuned to visual stimuli only and 27 cells (19%) were tuned to vestibular stimuli only. The distributions of direction preferences for rotation resembled those for translation overall (compare Figs. 6*A*,*B*, 4*A*,*B*). The distribution of visual azimuth preferences was significantly bimodal ($p < 0.001$, uniformity test; $p_{uni} = 0.02$, $p_{bi} = 0.97$, modality test), with a tendency for neurons to prefer pitch rather than roll rotations (Fig. 5*B*). Because of the limited size of the data set, none of the other distributions of azimuth or elevation preferences were significantly different from uniform ($p > 0.05$).

As for translation, rotation tuning was typically broad, with tuning width showing no significant dependence on azimuth preference (Fig. 6*C*,*D*). Compared with area MSTd, we found that the average tuning width in VIP was significantly narrower
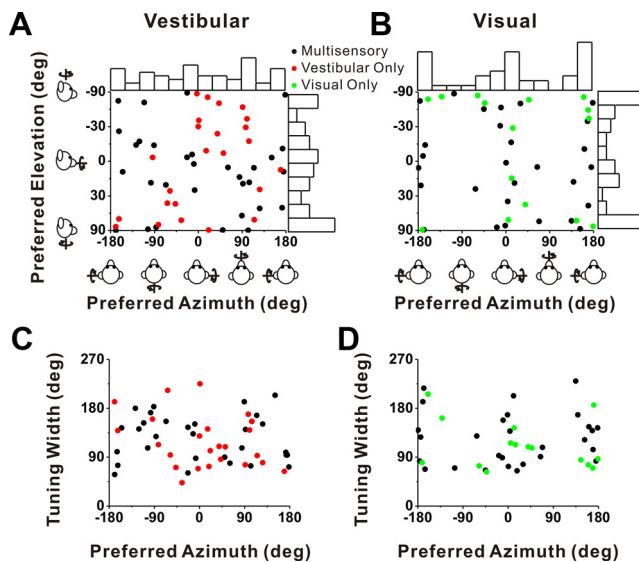
Figure 6. Summary of tuning properties of VIP neurons during rotation. *A*, *B*, Distribution of vestibular (*A*) and visual (*B*) 3D rotation preferences. Each data point in the scatter plot corresponds to the preferred azimuth (abscissa) and elevation (ordinate) of a single neuron with significant unimodal heading tuning (*A*, n = 53; *B*, n = 43). The format is as in Figure 4. *C*, *D*, Scatter plots of the tuning width at half-maximum of each cell versus preferred azimuth. The black symbols represent cells with significant unimodal tuning during both the vestibular (n = 31) and visual (n = 28) conditions. The red symbols represent cells with significant unimodal vestibular tuning only (n = 22). The green symbols represent cells with significant unimodal visual tuning only (n = 15).

Figure 7. Comparison of tuning preferences and direction selectivity between visual and vestibular responses of VIP neurons during rotation. *A*, Distribution of the absolute 3D difference in preferred direction (|Δ preferred direction|) between visual and vestibular responses (n = 23). The format is as in Figure 5. Only neurons with significant unimodal spatial tuning during both vestibular and visual conditions have been included. *B*, Scatter plot of the visual DDI as a function of the vestibular DDI. Black filled symbols, with significant tuning during both the vestibular and visual conditions (n = 36); red symbols, cells with significant tuning during the vestibular condition only (n = 27); green symbols, cells with significant tuning during the visual condition only (n = 24); open symbols, cells without significant tuning in either condition (n = 55). Dashed line, Unity-slope diagonal.

for both the visual and vestibular rotation conditions (p < 0.001, Wilcoxon's rank sum test).

Nearly all vestibular and visual rotation responses in VIP were classified as single-peaked (Table 1). As illustrated in Figure 7*A*, the distribution of the absolute difference in 3D preferred direction (|Δ preferred direction|) between visual and vestibular rotation responses of multisensory cells was significantly bimodal (p = 0.01, uniformity test; p_{uni} < 0.001; p_{bi} = 0.61, modality test). Thus, similar to translation responses, VIP neurons can show either congruent or opposite rotational preferences for visual and vestibular stimuli. As discussed further below, this property of VIP neurons differs from that seen in MSTd, where almost all
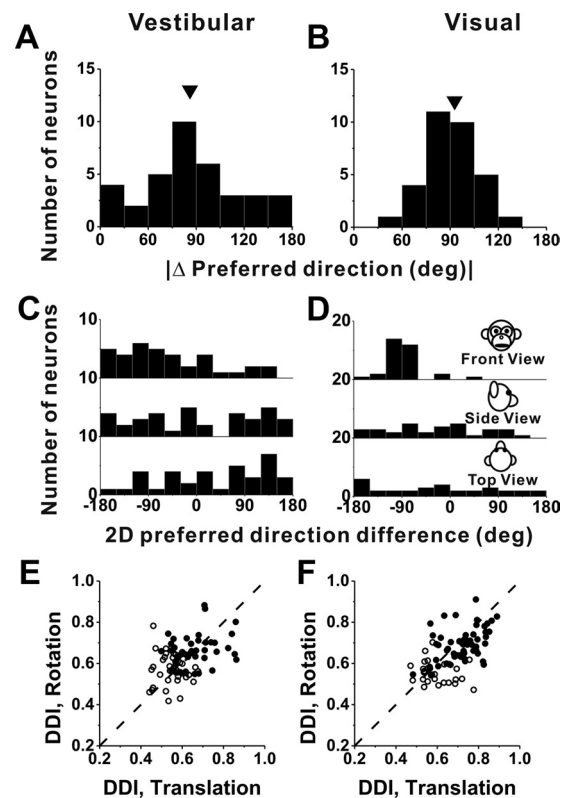

Figure 8. Summary of differences in direction preference and tuning strength between rotation and translation. *A*, *B*, Histograms of the absolute differences in 3D preferred direction (|Δ preferred direction|) between rotation and translation for the vestibular (n = 36) and visual (n = 32) conditions, respectively (calculated only for neurons with significant unimodal tuning in both conditions). The arrows illustrate mean values. *C*, *D*, Distributions of preferred direction differences projected onto each of the three cardinal planes: frontoparallel (front view), sagittal (side view), and horizontal (top view). *E*, *F*, Scatter plots of the rotation and translation DDI values for the vestibular and visual conditions, respectively. The filled symbols indicate cells with significant tuning for both rotation and translation (ANOVA, p < 0.01) (*E*, n = 45; *F*, n = 55); the open symbols denote cells without significant tuning for either one or both of the rotation and translation protocols (ANOVA, p > 0.01) (*E*, n = 37; *F*, n = 27). Dashed lines, Unity-slope diagonals.

neurons have opposite rotation preferences for visual and vestibular stimuli (Takahashi et al., 2007).

Visual and vestibular rotation responses in VIP had similar tuning strength, as illustrated in Figure 7*B* (Wilcoxon's matched-pairs test, p = 0.20; n = 142). As for translation tuning (Fig. 5*B*), the visual DDI of multisensory neurons (0.70 ± 0.01) was greater, on average, than the visual DDI of unisensory neurons (0.68 ± 0.01; p < 0.001, Wilcoxon's rank sum test) (Fig. 7*B*, compare black symbols, green symbols). A similar trend was seen for the vestibular DDI values, but the difference was not significant (p = 0.21, Wilcoxon's rank sum test) (Fig. 7*B*, compare black symbols, red symbols).

Comparison between translation and rotation responses
Since all cells tested with the rotation protocol were also tested with the translation protocol, a direct comparison between rotation and translation tuning is possible. The distribution of the absolute difference in 3D direction preference, |Δ preferred direction|, between vestibular rotation and vestibular translation conditions was not significantly different from uniform (p = 0.20, uniformity test), although there was some tendency for vestibular translation and rotation preferences to differ by ∼90° (Fig. 8*A*). That is, cells that prefer lateral translation (0, 180°) also
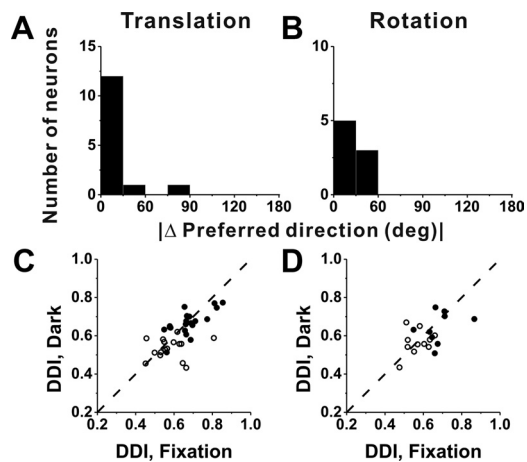
Figure 9. Comparison between VIP responses during fixation and in darkness. **A**, **B**, Distribution of the absolute difference in preferred direction for neurons with significant unimodal tuning during both fixation and in complete darkness; data are shown for translation (**A**) (n = 14) and rotation (**B**) (n = 8) separately. **C**, **D**, Scatter plot of DDI values for cells tested in both fixation and darkness conditions during translation (**C**) (n = 36) and rotation (**D**) (n = 20). Filled symbols, Cells with significant spatial tuning during both fixation and darkness (**C**, n = 20; **D**, n = 9). Open symbols, Cells without significant spatial tuning during either fixation or total darkness (**C**, n = 16; **D**, n = 11). Dashed lines, Unity-slope diagonals.

tend to prefer roll rotation (±90°), although the effect did not reach statistical significance.

The corresponding distribution of differences in direction preference between visual rotation and translation responses was significantly nonuniform ($p < 0.001$, uniformity test) and unimodal ($p_{uni} = 0.88$, modality test; mean, 90.0 ± 2.8°) (Fig. 8B). This relationship is not surprising when one considers that, at least for lateral–vertical optic flow (the preferred stimuli for many VIP cells) (Figs. 4B, 6B), visual translation and rotation preferences are typically linked by the two-dimensional visual motion selectivity of the cell. For example, a neuron that prefers leftward visual motion on the display screen will respond well to both a yaw rotation stimulus (azimuth, 0°; elevation, 90°) and a lateral (rightward) translation stimulus (azimuth, 0°; elevation, 0°). Note that these two stimulus directions are 90° apart on the sphere (Fig. 2A).

Since |Δ preferred direction| is computed as the smallest angle between a pair of preferred direction vectors in 3D within the interval of (0, 180°), it is not known whether the observed peak near 90° in Figure 8, A and B, is derived from a single mode at −90° or from two modes at +90 and −90°. To examine this, we also illustrate the differences between translation and rotation preferences in each cardinal plane: front view, side view, and top view (Fig. 8C,D) over the entire 360° range. For the visual condition, the distribution of 2D direction differences in the frontal plane is more revealing than those in the other two planes: data are tightly clustered around −90°, with no cells having direction differences of +90°. Thus, the data from the visual condition are mostly consistent with the idea that the preferred directions for translation and rotation are related through the 2D visual motion selectivity of VIP neurons. Note that the visual fixation target is head fixed in this study, such that both yaw/pitch rotations and lateral translations produce laminar optic flow in which all elements move in the same direction on the display screen. Although the speed of dot motion varies with distance for translation (but not rotation), the argument here depends only on the direction preferences for rotation and translation and not on speed selectivity.

Unlike in the visual stimulus condition, differences in direction preference between translation and rotation for the vestibular condition were not tightly distributed in any of the cardinal planes (Fig. 8C), a finding that is similar to that reported previously for MSTd (Takahashi et al., 2007), thalamus (Meng et al., 2007), and vestibular nuclei neurons (Dickman and Angelaki, 2002; Bryan and Angelaki, 2009). In contrast, vestibular preferences for rotation and translation in PIVC were spatially coordinated such that each cell preferring a given translation direction also responded maximally to rotations about an axis that was roughly perpendicular to the translation preference (Chen et al., 2010). For example, cells that preferred left–right translation also preferred roll rotation; and cells that preferred forward–backward translation tended to prefer pitch rotation. Thus, vestibular responses in VIP and MSTd differ from those in PIVC in that rotation/translation preferences are not as tightly aligned.

Tuning strength was not significantly different between vestibular translation (DDI, 0.62 ± 0.01 SE) and vestibular rotation (0.61 ± 0.01) conditions ($p = 0.27$, Wilcoxon's matched-pairs test), as illustrated in Figure 8E. By comparison, the visual rotation DDI (0.65 ± 0.01) was less than the visual translation DDI (0.68 ± 0.01) on average, and this difference was significant ($p = 0.001$, Wilcoxon's matched-pairs test) (Fig. 8F).

Fixation versus darkness

A subpopulation of VIP cells was also tested during vestibular translation (n = 36) and rotation (n = 20) in complete darkness (with the video projector turned off) (see Materials and Methods). As summarized in Figure 9, spatial tuning and response selectivity were similar whether the animal fixated a head-fixed target on the screen or was moved in darkness. This conclusion is based on two comparisons. First, for responses with significant unimodal spatial tuning under both conditions (translation, 14 of 36 cells; rotation, 8 of 20 cells), the distribution of the absolute difference in direction preference between fixation and darkness was narrow and biased strongly toward zero (median, 12.3° for translation and 40.1° for rotation) (Fig. 9A,B). With the exception of one cell, VIP neurons had similar (<60°) direction preferences in the fixation and darkness protocols. Second, tuning strength, as measured with the DDI, was not significantly different between fixation and darkness (Wilcoxon's matched-pairs test, $p = 0.09$ for translation; $p = 0.39$ for rotation), and DDI values for these two conditions were robustly correlated (Fig. 9C,D) ($r = 0.64$, $p < 0.001$ for translation; $r = 0.52$, $p = 0.02$ for rotation). These data suggest that responses in our vestibular condition mainly reflect sensory input from the vestibular apparatus, rather than either retinal slip or efferent eye movement signals. Note, however, that a somatosensory contribution to the vestibular tuning of VIP cells cannot be excluded.

Responses to combined visual/vestibular stimuli

To characterize the interaction between the two sensory modalities, a subset of neurons (50 cells for translation and 30 cells for rotation) was tested under three stimulus conditions (vestibular only, visual only, and combined stimulation) (see Materials and Methods). For congruent cells, the combined response had tuning that was similar to the single-cue responses, as shown for an example congruent cell in Figure 10A. In contrast, two different types of interactions were observed for opposite cells. For some cells, like the example in Figure 10B, the combined response was dominated by the vestibular tuning. For other cells, like that in

Figure 10C, the combined response was dominated by the visual tuning.

These observations are summarized in Figure 11A–D, which shows the distributions of |Δ preferred direction| between combined and vestibular conditions (Fig. 11A,B) or between combined and visual conditions (Fig. 11C,D). For congruent multisensory cells (black bars), |Δ preferred direction| was generally very small, as expected, indicating that the direction preference in the combined condition was similar to those in the visual and vestibular conditions. For opposite multisensory cells (gray bars), |Δ preferred direction| was broadly distributed, demonstrating that the direction preference of the combined response could be dominated by either the vestibular or visual preference. As expected, the combined tuning of unimodal cells (red/green bars) was similar to the unimodal responses. The fact that combined responses of multisensory VIP cells can be dominated by either the visual or vestibular tuning contrasts with previous results from area MSTd, for which combined responses under identical stimulus conditions (100% motion coherence) were generally dominated by the visual response (Gu et al., 2006; Takahashi et al., 2007; Morgan et al., 2008).

How does tuning strength of the combined response, as quantified by the DDI, compare with tuning strength of the single-cue responses? Across the population, combined responses to translation tend to have greater DDI values (0.71 ± 0.02 SE) than vestibular (0.63 ± 0.02 SE; $p < 0.001$, Wilcoxon's matched-pairs test) or visual responses (0.68 ± 0.02 SE; $p = 0.03$). Stronger tuning was also seen during combined responses to rotation (0.68 ± 0.02) than for vestibular (0.64 ± 0.02; $p = 0.02$, Wilcoxon's matched-pairs test) or visual (0.64 ± 0.02; $p = 0.11$) rotation responses, although the latter difference did not reach significance. Interestingly, visual-only and vestibular-only cells (Fig. 11E–H, open green and red symbols) also tend to show greater DDI (stronger tuning) during cue combination (translation, $p < 0.001$ for visual only and $p = 0.22$ for vestibular only; rotation, $p = 0.13$ for visual only and $p < 0.001$ for vestibular only; Wilcoxon's rank sum tests).

To explore further the relative contributions of vestibular and visual responses to the combined tuning of multisensory neurons, we fit the combined responses of each neuron with a weighted linear sum of the visual and vestibular responses (see Materials and Methods). The linear model generally provided very good fits to the combined responses. For translation, the median values of R^2 are 0.91 for VIP, 0.87 for VPS, and 0.92 for MSTd, respectively. For rotation, the median values of R^2 are 0.78 for VIP and 0.97 for MSTd. Note that this is consistent with findings of a previous study of MSTd neurons (Morgan et al., 2008), for which inclusion of a larger range of stimuli allowed comparison of linear and nonlinear models, with little explanatory power gained when including nonlinear terms. In addition, we only included cells with good fits of the linear model ($R^2 > 0.7$) in the following analyses. From these fits, we computed vestibular and visual gains, as well as the gain ratio (Fig.
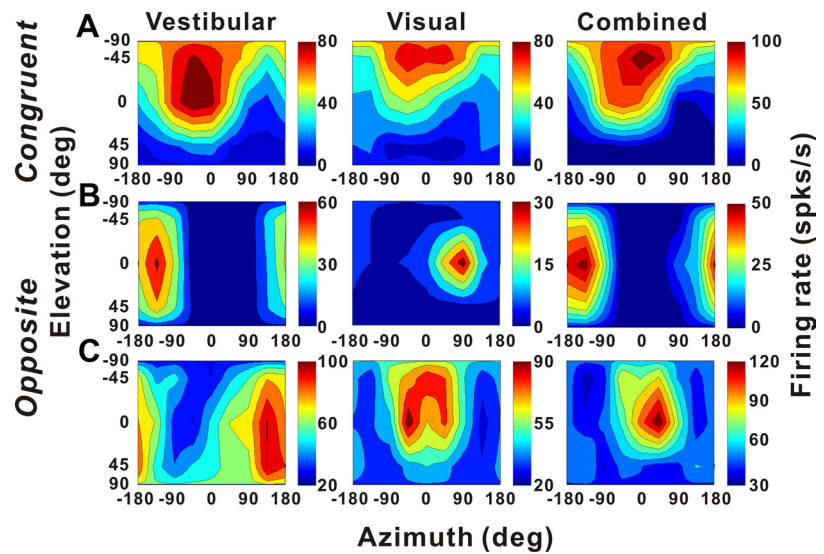
Figure 10. Examples of 3D translation tuning for three VIP neurons tested in the vestibular (left), visual (middle), and combined (right) conditions. The format is as in Figure 2. **A**, Tuning of a congruent multisensory neuron. Vestibular condition: Direction preference, [−40, −55°]; DDI, 0.85; visual condition: direction preference, [−32, −73°]; DDI, 0.84; combined condition: direction preference, [−31, −61°]; DDI, 0.88. **B**, Tuning of an opposite multisensory neuron for which the combined direction preference was dominated by the vestibular input. Vestibular condition: Direction preference, [−149, 0°]; DDI, 0.86; visual condition: direction preference, [88, −17°]; DDI, 0.78; combined condition: direction preference, [−159, −1°]; DDI, 0.83. **C**, Tuning of an opposite multisensory neuron for which the combined preference was dominated by the visual input. Vestibular condition: [136, 23°]; DDI, 0.55; visual condition: direction preference, [−6, −31°]; DDI, 0.63; combined condition: direction preference, [26, −19°]; DDI, 0.65.

12A,B). These gains describe the weighting of visual and vestibular inputs to the combined response. A gain ratio of 1 indicates that vestibular and visual inputs are equally weighted in the combined response; a gain ratio of <1 indicates that vestibular inputs contribute less than visual inputs; and a gain ratio of >1 means that vestibular inputs outweigh the visual inputs.

For translation, the mean gain ratio in VIP was 1.33 ± 0.79 (geometric mean, mean ± SE), with roughly one-half of the neurons having gain ratios of <1 and >1 (Fig. 12A, top). This value was significantly greater than that for MSTd (0.33 ± 0.29) (Fig. 12A, middle) ($p < 0.001$, Wilcoxon's rank sum test), but not significantly less than that for VPS (2.29 ± 1.52) (Fig. 12A, bottom) ($p = 0.14$). A similar tendency was observed for responses to rotation, with a mean gain ratio of 1.05 ± 1.08 for VIP and 0.41 ± 0.92 for MSTd, although this difference did not quite reach significance due to the smaller samples ($p = 0.06$, Wilcoxon's rank sum test). Note that we did not measure visual rotation responses for area VPS; hence no rotational gain ratios are presented for VPS. Overall, this analysis supports the findings of Figures 10 and 11, namely that vestibular cues contribute more strongly to combined responses in VIP than in MSTd, but less strongly than in VPS.

Unlike in MSTd (Takahashi et al., 2007), there was no dependence of the gain ratio on the relative strength of visual and vestibular responses in VIP (measured by the "visual-vestibular ratio") (see Materials and Methods) (Spearman's rank correlation, $p > 0.3$). That is, the relative tuning strength of the two single cues was not predictive of how the cues interact to determine the combined response in VIP.

Characterization of spatiotemporal dynamics
To assess the flow of visual and vestibular signals through a network of cortical areas involved in self-motion analysis, it is
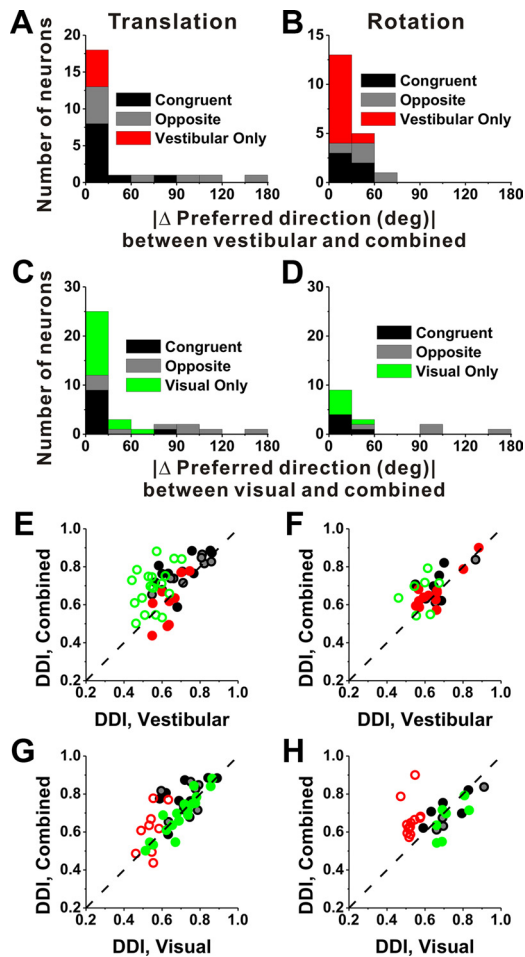
Figure 11. Summary of the differences in direction preference and comparison of tuning strength between the combined condition and each of the vestibular and visual conditions. *A*, *B*, Distributions of the absolute difference in 3D preferred direction (|∆ preferred direction|) between the combined condition and the vestibular condition, for responses obtained during translation (*A*) and rotation (*B*). *C*, *D*, Distributions of |∆ preferred direction| between the combined condition and the visual condition, for both translation (*C*) and rotation (*D*). *E–H*, Scatter plots comparing the combined DDI against either the vestibular or the visual DDI. Filled symbols, Cells for which both the combined and vestibular (*E*, *G*) or visual (*F*, *H*) tuning was significant (ANOVA, $p < 0.01$). Open symbols, for which either the combined and/or the vestibular/visual tuning was not significant (ANOVA, $p > 0.01$). Black symbols/bars, Multisensory congruent neurons (*E*, *G*, $n = 10$; *F*, *H*, $n = 5$); gray symbols/bars, multisensory opposite neurons (*E*, *G*, $n = 9$; *F*, *H*, $n = 4$); red symbols/bars, vestibular-only neurons (*E*, *G*, $n = 9$; *F*, *H*, $n = 12$); green symbols/bars, visual-only neurons (*E*, *G*, $n = 19$; *F*, *H*, $n = 7$). Dashed line, Unity-slope diagonal.
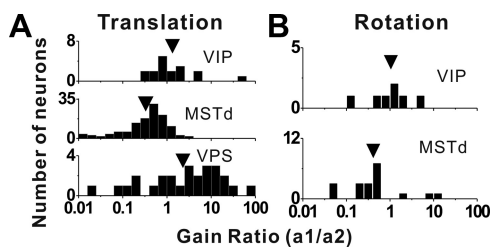
Figure 12. Distributions of the gain ratio, describing the relative weighting of the visual and vestibular contributions to the combined response for translation (*A*) and rotation (*B*). *A*, Top row, Data from area VIP ($n = 17$); middle row, data from area MSTd ($n = 125$); bottom row, data from area VPS ($n = 26$). The arrows illustrate geometric mean values. *B*, Top row, Data from VIP ($n = 7$); bottom row, data from area MSTd ($n = 19$). Only data with significant spatial tuning for both visual and vestibular stimuli and with good fits of the linear model ($R^2 > 0.7$) are included in this analysis.

valuable to examine the timing of responses. However, response timing has to be evaluated while accounting for potential variations in response dynamics (e.g., velocity vs acceleration coding) across areas. We have recently developed a model-based approach to this issue (Chen et al., 2011a), and we apply it here to visual/vestibular responses in areas VIP, MSTd, and VPS.

Three different models, reflecting coding of velocity (model V), velocity plus acceleration (model VA), or velocity plus acceleration plus position (model VAP), were fit to the spatiotemporal responses of each neuron (for details, see Materials and Methods). To reduce dimensionality, each model was fit to two-dimensional data from the horizontal, frontal, and median planes of the spherical stimulus space (Fig. 2*A*), provided that significant space–time structure was exhibited for each plane as described in Materials and Methods. We report results from the plane with the strongest response modulation for each neuron. In addition, we only quantified results from cells for which the goodness of fit of the model was high ($R^2 > 0.7$) for the best response plane. For the vestibular condition, 30 VIP neurons, 48 MSTd cells, and 24 VPS neurons met these criteria. For the visual condition, 76 cells from VIP, 119 from MSTd, and 14 from VPS met the criteria. Only 7 VIP neurons passed these criteria for the combined condition; hence we do not present fit results for that condition. Similarly, there was insufficient data from the rotation protocols for this analysis.

Figure 13 shows fits of models V, VA, and VAP to visual responses from an example neuron in area VIP. Responses from this neuron were equally well fit by model V (Fig. 13*B*), model VA (Fig. 13*C*), and model VAP (Fig. 13*D*), as illustrated by the fit residuals. Indeed, the fits of the three models (Fig. 13*E*, red, green and blue traces) are highly overlapping, consistent with a velocity weight, $w_v = 0.825$ for model VA and a position weight, $w_p = 0.031$ for model VAP. For this neuron, model V accounts for 93.4% of the variance in the data (93.5% for model VA; 93.6% for model VAP) and thus provides a good description of the spatiotemporal response profile. Models VA and VAP are not justified given the increase in the number of parameters ($p > 0.36$, sequential F test). For visual heading tuning, this pattern of results was observed for the majority of neurons in areas VIP and MSTd (Table 2, Visual response).

To summarize the relative strengths of velocity, acceleration, and position components in the neural responses, we computed the ratio of acceleration and velocity weights, w_a/w_v, as well as the position weight, w_p (see Materials and Methods). Cumulative distributions of these weights for each area are shown in Figure 14. For the vestibular stimulus condition, the acceleration to velocity weight ratio, w_a/w_v, was significantly larger in VIP (geometric mean ± SE, 1.59 ± 0.86) than in MSTd (0.70 ± 0.72 SE) ($p = 0.004$, Wilcoxon's rank sum test), but not significantly different from VPS where the sample size was rather small (1.15 ± 1.00 SE) ($p = 0.44$). The position weight, w_p, was very small for VIP (0.05 ± 0.55 SE) and VPS (0.07 ± 0.68 SE), but significantly greater for MSTd (0.14 ± 0.41 SE) than either VIP or VPS ($p < 0.001$, Wilcoxon's rank sum tests). Thus, for vestibular responses to translation, activity in VIP (and VPS) reflects fairly balanced contributions of velocity and acceleration, with little position component. In contrast, vestibular responses in MSTd are more dominated by stimulus velocity, with a modest but more substantial contribution from position.

For the visual stimulus condition, the ratio of acceleration and velocity weights (w_a/w_v) was relatively low (VIP, 0.22 ± 0.32 SE; MSTd, 0.17 ± 0.42 SE; VPS, 0.16 ± 0.66 SE), reflecting
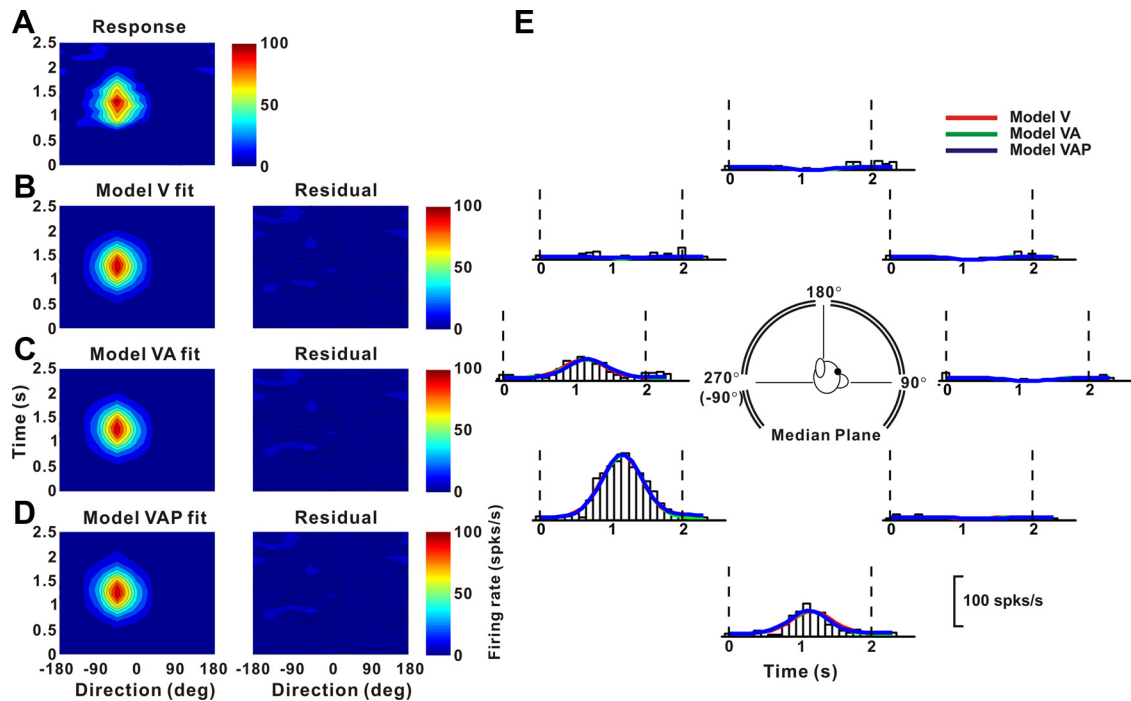
Figure 13. Example fits of velocity (model V), velocity plus acceleration (model VA), and velocity plus acceleration plus position (model VAP) models to the spatiotemporal visual responses of a VIP neuron. *A*, Direction–time plot showing how direction tuning evolves over the time course of the response (spatial and temporal resolution, 45° and 100 ms, respectively). *B–D*, Model fits (left) and response residuals (right). For model V (*B*), $t_0 = 1.042$ s; $\theta = -43.2°$, $R^2 = 0.934$. For model VA (*C*), $t_0 = 1.148$ s; $\theta_0 = -43.3°$, $\Delta\theta_{va} = 5.9°$, $w_v = 0.825$, $R^2 = 0.935$. For model VAP (*D*), $t_0 = 1.149$ s; $\theta_0 = -43.3°$, $\Delta\theta_{va} = 5.5°$, $w_v = 0.81$, $\Delta\theta_{vp} = 5.5°$, $w_p = 0.031$, $R^2 = 0.936$. *E*, Response PSTHs (open bars) for the eight directions of motion in the median plane (see inset) along with superimposed fits of model V (red), model VA (green), and model VAP (blue). The vertical dashed lines mark the 2 s duration of the stimulus.

Table 2. Summary of best fitting models

	VIP [n (%)]	MSTd [n (%)]	VPS [n (%)]
Vestibular response			
Velocity only (model V)	0 (0)	1 (2)	3 (12)
Velocity plus acceleration (model VA)	17 (57)	11 (23)	12 (50)
Velocity plus acceleration plus position (model VAP)	13 (43)	36 (75)	9 (38)
All	30	48	24
Visual response			
Velocity only (model V)	44 (58)	80 (67)	5 (36)
Velocity plus acceleration (model VA)	25 (33)	20 (17)	5 (36)
Velocity plus acceleration plus position (model VAP)	7 (9)	19 (16)	4 (28)
All	76	119	14

Values shown represent the number (and percentage) of neurons in areas VIP, MSTd, and VPS that were better fit by model V, model VA, or model VAP according to the sequential *F* test (see Materials and Methods). Only cells from best fit plane with goodness of fit $R^2 > 0.7$ were included here.

dominance of stimulus velocity, and there were no significant differences across areas ($p > 0.22$, Wilcoxon's rank sum tests). The position weight was also low in VIP (0.04 ± 0.31 SE) and MSTd (0.05 ± 0.31 SE), but was slightly greater in area VPS (0.10 ± 0.88 SE) ($p = 0.009$, Wilcoxon's rank sum test). Thus, visual heading responses were largely dominated by velocity in all three areas.

In addition to estimating the relative contributions of velocity, acceleration, and position signals to neural responses, the model-fitting analysis also allowed us to compute the overall latency of the response (parameter t_0 in Eq. 6), independent of the specific mixture of temporal response components needed to fit the response of each neuron. Figure 15 shows distributions of response delays, from the best-fitting model, for areas VIP, MSTd, and VPS. For the vestibular condition, response latency for VIP neurons (mean $\pm$ SE, 112 ± 26 ms) was significantly earlier than for
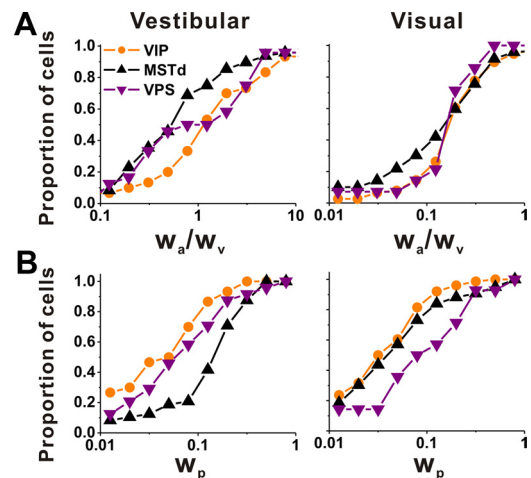
Figure 14. Population comparison of parameters of spatiotemporal model fits among VIP, MSTd, and VPS. *A*, Cumulative distributions of the ratio of acceleration to velocity weights (w_a/w_v) from model VA for the vestibular (left, VIP, $n = 30$; MSTd, $n = 48$; VPS, $n = 24$) and visual (right, VIP, $n = 76$; MSTd, $n = 119$; VPS, $n = 14$) conditions. *B*, Cumulative distributions of the position weight, w_p, from model VAP for the vestibular (left, VIP, $n = 30$; MSTd, $n = 48$; VPS, $n = 24$) and visual (right, VIP, $n = 76$; MSTd, $n = 119$; VPS, $n = 14$) conditions. Only data with good fits ($R^2 > 0.7$) were included here.

MSTd neurons (193 ± 26 ms) ($p = 0.007$, Wilcoxon's rank sum test), and not significantly different from that for VPS cells (66 ± 41 ms) ($p = 0.28$). For the visual condition, mean response latencies for the three areas did not differ significantly from each other (VIP, 48 ± 18 ms; MSTd, 77 ± 19 ms; VPS, 26 ± 73 ms; $p > 0.16$, Wilcoxon's rank sum tests). Thus, timing of vestibular responses in VIP was faster than in MSTd, whereas timing of optic flow responses did not differ significantly among areas.
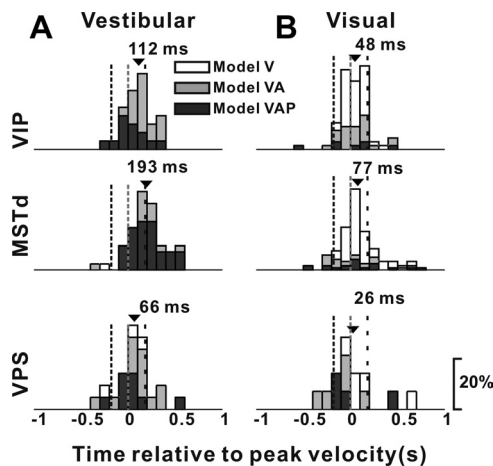
Figure 15. Distributions of response latency, derived from model fits, for neurons in VIP (top row), MSTd (middle row), and VPS (bottom row), as tested under the vestibular (**A**, VIP, $n = 30$; MSTd, $n = 48$; VPS, $n = 24$) and visual (**B**, VIP, $n = 76$; MSTd, $n = 119$; VPS, $n = 14$) conditions. Open bars, Cells better fit with model V; gray bars, cells better fit with model VA; black bars, cells better fit with model VAP. The arrows indicate mean values. The vertical dashed lines indicate the times of peak acceleration/deceleration and peak velocity of the stimulus.
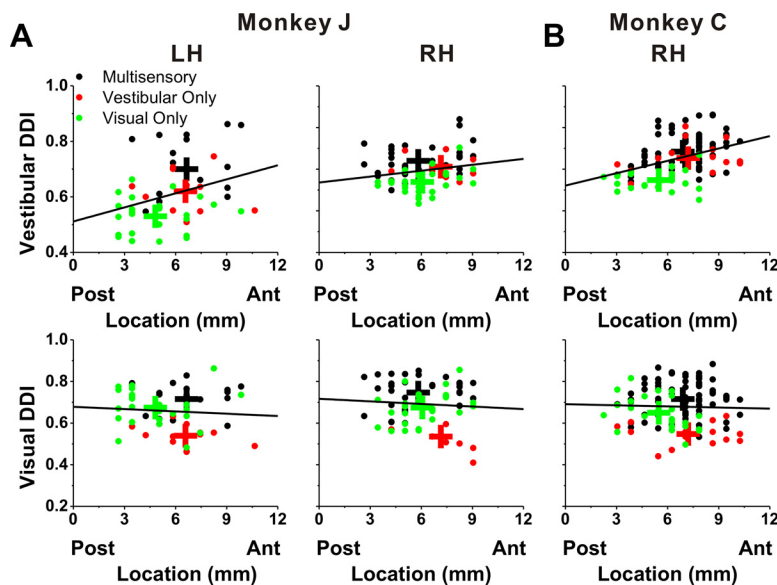
Figure 16. Relationship between tuning strength, as measured by DDI, and cell location within the intraparietal sulcus. DDI values for the vestibular condition (top row) and the visual condition (bottom row) are plotted as a function of the anterior–posterior stereotaxic coordinate of each electrode penetration (shown in millimeters, with 0 mm corresponding to AP-0). **A**, Data recorded from the left hemisphere (LH) of monkey J (left column, $n = 53$) and the right hemisphere (RH) of monkey J (right column, $n = 70$). **B**, Data obtained from the right hemisphere of monkey C ($n = 112$). Black symbols, Multisensory neurons; red symbols, vestibular-only neurons; green symbols, visual-only neurons. The solid lines illustrate the best-fit lines from linear regressions. Note that only responsive cells are included in these scatter plots. "+" marks the mean location for each group of cells.

Selectivity as a function of location within the intraparietal sulcus

Visual-only, vestibular-only, and multisensory cells were encountered throughout the anterior–posterior extent of VIP (Fig. 16). The relationship between DDI and anterior–posterior coordinates was examined for both the right and left hemispheres of monkey J (Fig. 16A) and the right hemisphere of monkey C (Fig. 16B), as all three were extensively explored. There was a significant correlation between the DDI for vestibular translation and the anterior–posterior coordinates (monkey J's left hemisphere:

$r = 0.32$, $p = 0.02$; monkey J's right hemisphere: $r = 0.20$, $p = 0.09$; monkey C's right hemisphere: $r = 0.40$, $p < 0.001$). No such relationship was seen for the visual translation condition ($p > 0.5$) (Fig. 16A,B, bottom panels). For rotation (likely due to the small sample of neurons), we did not see any significant correlations between DDI and the anterior–posterior location of electrode penetrations, and this was true for both the vestibular and visual conditions.

As shown by the different colors in Figure 16, multisensory neurons and unisensory neurons were intermingled in VIP and were found across the entire anterior–posterior extent in these animals. Similarly, we found all of these cell types independent of the medial–lateral location of the electrode penetration, and within the medial bank, lateral bank, and tip of the intraparietal sulcus.

Discussion

Using an experimental protocol identical with that used previously in areas MSTd (Gu et al., 2006; Takahashi et al., 2007) and VPS (Chen et al., 2011b), we characterized the tuning of VIP neurons in response to 3D translational and rotational movements that were defined by physical body motion alone, optic flow alone, or both cues together. There were both similarities and differences in the response properties of neurons in VIP versus MSTd and VPS. We next discuss these findings, first focusing on the visual response properties of VIP neurons.

Location and visual properties of VIP

Maunsell and van Essen (1983) first showed that projections from area MT terminate in the depths of the intraparietal sulcus, as later confirmed by Ungerleider and Desimone (1986). Colby et al. (1993) reported that neurons in a portion of this projection zone, mainly confined to the fundus of the intraparietal sulcus but extending onto both the lateral and medial banks, had prominent visual direction selectivity. This smaller, physiologically defined region in the depths of the sulcus was named VIP. At its maximum extent, VIP typically extends over 5 mm mediolaterally and 8 mm rostrocaudally (Colby et al., 1993). In two animals, we explored the full extent of VIP (Figs. 1, 16). Visual and vestibular responses were encountered throughout this area. Although no particular topographic organization was obvious, vestibular tuning was strongest in the anterior half of VIP (Fig. 16). Duhamel et al. (1998) also reported intermingling of visual-only and multisensory visual/tactile-responsive neurons and no systematic progression of receptive field locations within this area.

Unlike in MSTd, where nearly all (>95%) neurons are tuned for the direction of translation or rotation defined by optic flow, only 69 and 42% of VIP cells were tuned for translational and rotational optic flow, respectively. These values are in general agreement with the study by Bremmer et al. (2002b), who reported that ~75% of VIP neurons were responsive to optic flow. Our finding of a bimodal distribution

of visual azimuth preferences, showing that the least-preferred optic-flow stimuli involved translation along or rotation about the forward–backward axis (Figs. 4 *B*, 6 *B*), appears at odds with previous studies reporting that twice as many VIP cells preferred expansion and rotation over frontoparallel stimuli (Schaafsma and Duysens, 1996; Bremmer et al., 2002b). However, several differences in the visual stimuli used (e.g., wall vs cloud of dots, binocular disparity vs no disparity) and the cell selection criteria make direct comparisons difficult. Our findings show that, although VIP cells often respond to complex optic flow patterns like expansion/contraction and clockwise/counterclockwise rotation, their maximum responses are generally found along axes in the frontoparallel plane. The finding of underrepresented forward–backward motion directions in VIP, which is similar to our previous findings in MSTd (Gu et al., 2006, 2010), may provide a functional benefit. Given the broad spatial tuning of most VIP neurons (Figs. 4*C,D*, 6*C,D*) (but see Zhang and Britten, 2010), the bias toward lateral motion preferences may facilitate precise direction discrimination around the forward–backward axis. Indeed, population activity in MSTd, like human and monkey perception, is most sensitive for discriminating small changes in heading direction around straight forward rather than around lateral movement directions (Gu et al., 2010).

Previous studies have also emphasized that VIP neurons prefer expansion over contraction in optic flow stimuli. Schlack et al. (2002) reported that the dominance of expansion- over contraction-preferring neurons was particularly strong for visual/vestibular multisensory cells in VIP and was reduced when all visually responsive neurons were considered regardless of vestibular responses. Such a difference was not evident in our data (Fig. 4*B*).

Vestibular rotation and translation tuning in VIP

We found that nearly one-half of VIP neurons were tuned during 3D vestibular stimulation, similar to previous reports for area MSTd (Gu et al., 2006; Takahashi et al., 2007; Chowdhury et al., 2009). Our proportion of selective neurons is lower than the proportion of cells responding to forward or backward translation on a parallel swing (78%) (Schlack et al., 2002), but comparable with the proportion of cells that modulated during yaw rotation (approximately one-third) (Bremmer et al., 2002b). We recorded from any well isolated VIP cell, including any neuron that was either spontaneously active or driven by a full-field flickering random-dot stimulus. We found that vestibular tuning was stronger in anterior portions of VIP, also in agreement with the results of Bremmer et al. (2002b). Visual tuning did not vary with anterior–posterior location along the intraparietal sulcus.

In response to translation, both congruent and opposite neurons were encountered in VIP with roughly equal frequencies, similar to MSTd (Gu et al., 2006, 2010). Schlack et al. (2002) also reported both congruent and opposite selectivities during forward–backward translation. In a different study, however, visual and vestibular preferences for yaw rotation were reported to be consistently incongruent (Bremmer et al., 2002b). In contrast, we found that multisensory cells in VIP can be either congruent or opposite in response to rotations. The presence of congruent rotation responses in VIP is one of the main differences we encountered between VIP and MSTd; in MSTd, virtually all multisensory rotation responses had near-opposite visual and vestibular preferences (Takahashi et al., 2007; Chowdhury et al., 2009).

Multisensory responses in VIP— comparison with MSTd and VPS

Area VIP is widely acknowledged as a multisensory area (Colby et al., 1993; Duhamel et al., 1998; Bremmer et al., 2002b; Schlack et al., 2002, 2005; Schroeder and Foxe, 2002). Most studies of VIP have examined multisensory convergence of visual and tactile stimuli, for which ~70% of VIP neurons have multimodal responses (Duhamel et al., 1998; Avillac et al., 2005, 2007) with visual and tactile receptive fields that are typically aligned spatially. By contrast, visual/vestibular integration appears to be somewhat less frequent in VIP, although sampling biases cannot be excluded.

Visual/vestibular multisensory response properties differ between VIP, MSTd, and VPS in three main ways. First, when both cues are presented together (at 100% motion coherence), responses of VIP neurons are dominated either by visual or vestibular tuning, whereas visual responses generally dominate in MSTd and vestibular responses dominate in VPS. The balance of visual and vestibular influences in VIP is similar to that seen in MSTd when motion coherence is reduced (Morgan et al., 2008), consistent with stronger effective vestibular inputs to VIP. Second, congruent VIP cells are found during both translation and rotation stimulation. In contrast, all rotation-sensitive multisensory cells in MSTd have opposite tuning preferences (Takahashi et al., 2007; Chowdhury et al., 2009), and most VPS neurons have opposite preferences for visual and vestibular translation. Based on these properties, we previously speculated that rotation signals in MSTd and translation signals in VPS cannot be used for multisensory cue integration (Takahashi et al., 2007; Chen et al., 2011b). However, the presence of congruent visual/vestibular rotation responses in VIP may be consistent with roles in perception of both self-translation (heading) and self-rotation. A third difference between VIP, MSTd, and VPS involves visual/vestibular tuning strength. Optic flow tuning in VIP was weaker than in MSTd but stronger than in VPS. In contrast, vestibular translation tuning was slightly stronger and peaked earlier in VIP than in MSTd (Chen et al., 2011a) but was weaker than in VPS. These findings suggest that VIP is proximal to MSTd but distal to VPS in terms of vestibular signal processing.

Since we tested VIP, MSTd, and VPS neurons under identical conditions, our analyses allow quantitative comparisons across areas. However, these comparisons have only been made in the context of a passive fixation task, in which the visual and vestibular stimuli are not directly relevant to behavior. It is possible that other differences among areas could emerge in behavioral tasks that require animals to integrate cues and make judgments of self-motion. We have undertaken such studies in area MSTd (Gu et al., 2007, 2008, 2010; Fetsch et al., 2009), and similar studies are ongoing in VIP.

The functional significance of visual/vestibular integration in VIP remains unclear, although several proposals have been made in the literature. VIP has been implicated in planning and visual guidance of movements in near extrapersonal space (Colby et al., 1993; Bremmer et al., 2002a), in construction of a multisensory, head-centered representation of near extrapersonal space, or in planning defensive movements (Cooke and Graziano, 2003). Indeed, VIP is connected with F4, a region of inferior premotor cortex implicated in the control of head and mouth movements (Matelli et al., 1994; Lewis and Van Essen, 1996). Finally, others view VIP, like MSTd, as involved in visual/vestibular integration for perception of self-

motion (Bremmer et al., 2002a,b; Cook and Maunsell, 2002a,b; Zhang and Britten, 2010). Which of these functions are truly assignable to VIP remains to be determined.

References

Angelaki DE, Cullen KE (2008) Vestibular system: the many facets of a multimodal sense. Annu Rev Neurosci 31:125–150.

Anzai A, Peng X, Van Essen DC (2007) Neurons in monkey visual area V2 encode combinations of orientations. Nat Neurosci 10:1313–1321.

Avillac M, Denève S, Olivier E, Pouget A, Duhamel JR (2005) Reference frames for representing visual and tactile locations in parietal cortex. Nat Neurosci 8:941–949.

Avillac M, Ben Hamed S, Duhamel JR (2007) Multisensory integration in the ventral intraparietal area of the macaque monkey. J Neurosci 27:1922–1932.

Baizer JS, Ungerleider LG, Desimone R (1991) Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. J Neurosci 11:168–190.

Boussaoud D, Ungerleider LG, Desimone R (1990) Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. J Comp Neurol 296:462–495.

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.

Bremmer F, Graf W, Ben Hamed S, Duhamel JR (1999) Eye position encoding in the macaque ventral intraparietal area (VIP). Neuroreport 10:873–878.

Bremmer F, Duhamel JR, Ben Hamed S, Graf W (2002a) Heading encoding in the macaque ventral intraparietal area (VIP). Eur J Neurosci 16:1554–1568.

Bremmer F, Klam F, Duhamel JR, Ben Hamed S, Graf W (2002b) Visual-vestibular interactive responses in the macaque ventral intraparietal area (VIP). Eur J Neurosci 16:1569–1586.

Britten KH (2008) Mechanisms of self-motion perception. Annu Rev Neurosci 31:389–410.

Britten KH, van Wezel RJ (1998) Electrical microstimulation of cortical area MST biases heading perception in monkeys. Nat Neurosci 1:59–63.

Bryan AS, Angelaki DE (2009) Optokinetic and vestibular responsiveness in the macaque rostral vestibular and fastigial nuclei. J Neurophysiol 101:714–720.

Chen A, Henry E, DeAngelis GC, Angelaki DE (2007) Comparison of responses to three-dimensional rotation and translation in the ventral intraparietal (VIP) and medial superior temporal (MST) areas of rhesus monkey. Soc Neurosci Abstr 33:715–719.

Chen A, DeAngelis GC, Angelaki DE (2010) Macaque parieto-insular vestibular cortex: responses to self-motion and optic flow. J Neurosci 30:3022–3042.

Chen A, DeAngelis GC, Angelaki DE (2011a) A comparison of vestibular spatiotemporal tuning in macaque parietoinsular vestibular cortex, ventral intraparietal area, and medial superior temporal area. J Neurosci 31:3082–3094.

Chen A, DeAngelis GC, Angelaki DE (2011b) Convergence of vestibular and visual self-motion signals in an area of the posterior sylvian fissure. J Neurosci 31:11617–11627.

Chowdhury SA, Takahashi K, DeAngelis GC, Angelaki DE (2009) Does the middle temporal area carry vestibular signals related to self-motion? J Neurosci 29:12020–12030.

Colby CL, Duhamel JR, Goldberg ME (1993) Ventral intraparietal area of the macaque: anatomic location and visual response properties. J Neurophysiol 69:902–914.

Cook EP, Maunsell JH (2002a) Attentional modulation of behavioral performance and neuronal responses in middle temporal and ventral intraparietal areas of macaque monkey. J Neurosci 22:1994–2004.

Cook EP, Maunsell JH (2002b) Dynamics of neuronal responses in macaque MT and VIP during motion detection. Nat Neurosci 5:985–994.

Cooke DF, Graziano MS (2003) Defensive movements evoked by air puff in monkeys. J Neurophysiol 90:3317–3329.

Dickman JD, Angelaki DE (2002) Vestibular convergence patterns in

vestibular nuclei neurons of alert primates. J Neurophysiol 88: 3518–3533.

Duffy CJ (1998) MST neurons respond to optic flow and translational movement. J Neurophysiol 80:1816–1827.

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.

Felleman DJ, Van Essen DC (1991) Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1:1–47.

Fetsch CR, Wang S, Gu Y, DeAngelis GC, Angelaki DE (2007) 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.

Fetsch CR, Turner AH, DeAngelis GC, Angelaki DE (2009) Dynamic reweighting of visual and vestibular cues during self-motion perception. J Neurosci 29:15601–15612.

Gu Y, Watkins PV, Angelaki DE, DeAngelis GC (2006) Visual and nonvisual contributions to three-dimensional heading selectivity in the medial superior temporal area. J Neurosci 26:73–85.

Gu Y, DeAngelis GC, Angelaki DE (2007) A functional link between area MSTd and heading perception based on vestibular signals. Nat Neurosci 10:1038–1047.

Gu Y, Angelaki DE, DeAngelis GC (2008) Neural correlates of multisensory cue integration in macaque MSTd. Nat Neurosci 11:1201–1210.

Gu Y, Fetsch CR, Adeyemo B, DeAngelis GC, Angelaki DE (2010) Decoding of MSTd population activity accounts for variations in the precision of heading perception. Neuron 66:596–609.

Klam F, Graf W (2003) Vestibular response kinematics in posterior parietal cortex neurons of macaque monkeys. Eur J Neurosci 18:995–1010.

Lewis JW, Van Essen DC (1996) Connectioins of visual area VIP with somatosensory and motor areas of the macaque monkey. Soc Neurosci Abstr 22:160.164.

Lewis JW, Van Essen DC (2000) Corticocortical connections of visual, sensorimotor, and multimodal processing areas in the parietal lobe of the macaque monkey. J Comp Neurol 428:112–137.

Liu S, Angelaki DE (2009) Vestibular signals in macaque extrastriate visual cortex are functionally appropriate for heading perception. J Neurosci 29:8936–8945.

Maciokas JB, Britten KH (2010) Extrastriate area MST and parietal area VIP similarly represent forward headings. J Neurophysiol 104:239–247.

Matelli M, Luppino G, Murata A, Sakata H (1994) Independent anatomical circuits for reaching and grasping linking the inferior parietal sulcus and inferior area 6 in macaque monkey. Soc Neurosci Abstr 20:404.404.

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.

Meng H, May PJ, Dickman JD, Angelaki DE (2007) Vestibular signals in primate thalamus: properties and origins. J Neurosci 27:13590–13602.

Morgan ML, DeAngelis GC, Angelaki DE (2008) Multisensory integration in macaque visual cortex depends on cue reliability. Neuron 59:662–673.

Nguyenkim JD, DeAngelis GC (2003) Disparity-based coding of three-dimensional surface orientation by macaque middle temporal neurons. J Neurosci 23:7117–7128.

Page WK, Duffy CJ (2003) Heading representation in MST: sensory interactions and population encoding. J Neurophysiol 89:1994–2013.

Schaafsma SJ, Duysens J (1996) Neurons in the ventral intraparietal area of awake macaque monkey closely resemble neurons in the dorsal part of the medial superior temporal area in their responses to optic flow patterns. J Neurophysiol 76:4056–4068.

Schlack A, Hoffmann KP, Bremmer F (2002) Interaction of linear vestibular and visual stimulation in the macaque ventral intraparietal area (VIP). Eur J Neurosci 16:1877–1886.

Schlack A, Sterbing-D'Angelo SJ, Hartung K, Hoffmann KP, Bremmer F (2005) Multisensory space representations in the macaque ventral intraparietal area. J Neurosci 25:4616–4625.

Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. Brain Res Cogn Brain Res 14:187–198.

Seltzer B, Pandya DN (1986) Posterior parietal projections to the intraparietal sulcus of the rhesus monkey. Exp Brain Res 62:459–469.

Silverman BW (1981) Using kernel density estimates to investigate multimodality. J R Stat Soc 43:97–99.

Takahashi K, Gu Y, May PJ, Newlands SD, DeAngelis GC, Angelaki DE (2007) Multimodal coding of three-dimensional rotation and translation in area MSTd: comparison of visual and vestibular selectivity. J Neurosci 27:9742–9756.

Ungerleider LG, Desimone R (1986) Cortical connections of visual area MT in the macaque. J Comp Neurol 248:190–222.

Van Essen DC, Lewis JW, Drury HA, Hadjikhani N, Tootell RB, Bakircioglu M, Miller MI (2001) Mapping visual cortex in monkeys and humans using surface-based atlases. Vision Res 41:1359–1378.

Warren WH (2004) Optic flow. In: The visual neuroscience (Chalupa LM, Werner JS, eds), pp 1247–1259. Cambridge, MA: MIT.

Zhang T, Britten KH (2004) Clustering of selectivity for optic flow in the ventral intraparietal area. Neuroreport 15:1941–1945.

Zhang T, Britten KH (2010) The responses of VIP neurons are sufficiently sensitive to support heading judgments. J Neurophysiol 103:1865–1873.

Zhang T, Heuer HW, Britten KH (2004) Parietal area VIP neuronal responses to heading stimuli are encoded in head-centered coordinates. Neuron 42:993–1001.