units (SUs) to the responses of multiunit (MU) activity recorded from the same microelectrode (after removing SU spikes from the MU activity). We analyzed whether neurons in MSTd are clustered in two ways. First, we measured the similarity between the 3D translation and rotation tuning of SU and MU responses. Second, we measured the similarity between the directional and spatial RF structure of SU and MU responses. The data presented here were collected as part of previous studies (Gu et al. 2006; Takahashi et al. 2007) in which SU responses were examined in detail. Due to the length of the experimental protocols in those studies, we generally did not collect data from multiple recording sites along penetrations through MSTd. Hence our analyses are limited to a comparison of SU and MU responses at the same recording site and we do not directly address variations in tuning as a function of distance along electrode penetrations.

Clustering of 3D translation and rotation tuning in MSTd

In the translation protocol (see METHODS), each MSTd neuron was tested with 26 directions of translation, consisting of all combinations of azimuth and elevation separated by 45° on a sphere (Fig. 1, *A* and *B*). Figure 2*A* shows an example of 3D translation tuning for simultaneously recorded SU and MU activity in MSTd. The data are shown as contour maps in which mean firing rate (represented by color) is plotted as a function of azimuth (abscissa) and elevation (ordinate) (see also Gu et al. 2006). Data in the *top row* show responses obtained in the vestibular stimulus condition and data in the *bottom row* are from the visual condition. Note that the MU responses are about twofold larger than the SU responses. In the vestibular condition, the SU shows clear spatial tuning for translation, with a preferred direction at 190° azimuth and −1° elevation. A nearly opposite translation preference was seen in the visual condition for this SU, with the direction preference occurring at 47° azimuth and 9° elevation. This pattern of results is typical of an "opposite" cell, as described previously (Gu et al. 2006). The MU activity recorded simultaneously with this SU (Fig. 2*A*, *right column*) shows similar tuning for translational motion with a preferred direction of (181°, 0°) for the vestibular condition and (9°, −33°) for the visual condition. This suggests that nearby neurons in MSTd have similar direction preferences.

The similarity in tuning between MU and SU responses cannot simply be due to the same single unit contributing to both signals. To avoid this confound, we have excluded SU spikes from the MU signal (see METHODS), so that the MU response reflects the combined activity of several other nearby SUs. This was verified by cross-correlation analysis, as shown in Fig. 2*B*. Before SU spikes were removed, there was a sharp peak in the cross-correlogram centered around 0 ms. After SU spikes were removed, the cross-correlogram was relatively flat, indicating that spikes from the single cell were effectively excluded from the MU signal. Thus the observed similarity in SU and MU tuning for the example in Fig. 2*A* is not attributable to a common source of spikes. All MU data reported herein had the corresponding SU spikes removed.

For the translation protocol, this analysis was performed on a total of 285 MSTd neurons for the vestibular condition and 270 MSTd neurons for the visual condition. All SU/MU pairs were recorded simultaneously from a single microelectrode. Table 1 summarizes the proportions of SU and MU responses with significant spatial tuning for translation. In the vestibular condition, 56% (161/285) of SUs and 32% (92/285) of MUs had significant spatial tuning (ANOVA, $P < 0.05$, Table 1). In contrast, 97% (261/270) of SUs and 82% (222/270) of MUs were significantly tuned in the visual translation condition. When the SU was significantly tuned, the MU was also significantly tuned in 45% of cases for the vestibular condition, compared with 84% of cases in the visual condition. When the MU was significantly tuned, the SU was also significantly tuned in 78% of cases for the vestibular condition, compared with 98% cases in the visual condition (see Table 1). Both SU and MU selectivities for translation were less prevalent in the vestibular condition.

For the rotation protocol, we analyzed data from a total of 81 and 66 SU/MU pairs for the vestibular and visual conditions, respectively. An example data set is shown in Fig. 3. In the vestibular rotation condition, the SU was spatially tuned with a direction preference at 150° azimuth and −73° elevation, whereas the MU preferred 88° azimuth and −69° elevation. In the visual condition, the peak SU response occurred at 350° azimuth and 69° elevation and the MU response peaked at 10° azimuth and 86° elevation. Thus visual and vestibular rotation preferences were similar for SU and MU activity for this example neuron. The corresponding cross-correlograms between SU and MU responses are shown in Fig. 3*B*. Again, it can be seen that coupling between SU and MU signals was effectively removed by our procedure.

As shown in Table 2, 93% (75/81) of SUs and 70% (57/81) of MUs were significantly tuned for vestibular rotation, compared with 99% (65/66) of SUs and 88% (58/66) of MUs for visual rotation (ANOVA, $P < 0.05$). When MU tuning was significant, SU tuning was also significant for 98% of SUs in both the vestibular and visual conditions. When SU tuning was significant, MU tuning was also significant for 75% of cases for the vestibular and 88% of cases for the visual conditions. Thus the incidence of significant tuning in MU activity was higher for vestibular rotation than that for vestibular translation. This may reflect tighter clustering of vestibular rotation responses or, simply, the fact that vestibular rotation tuning is generally stronger than vestibular translation tuning for SUs (Takahashi et al. 2007). Vestibular rotation responses may also reflect a contribution of retinal slip of the faint visual background since these rotation responses were found to be slightly weaker during viewing in total darkness (Takahashi et al. 2007).

TABLE 1. *Percentage of SU and MU with significant spatial tuning for translation (ANOVA, $P \leq 0.05$)*

	MU	
SU	$P \leq 0.05$	$P > 0.05$
Vestibular (n = 285)		
$P \leq 0.05$	72/285 (25%)	89/285 (31%)
$P > 0.05$	20/285 (7%)	104/285 (36%)
Visual (n = 270)		
$P \leq 0.05$	219/270 (81%)	42/270 (16%)
$P > 0.05$	3/270 (1%)	6/270 (2%)

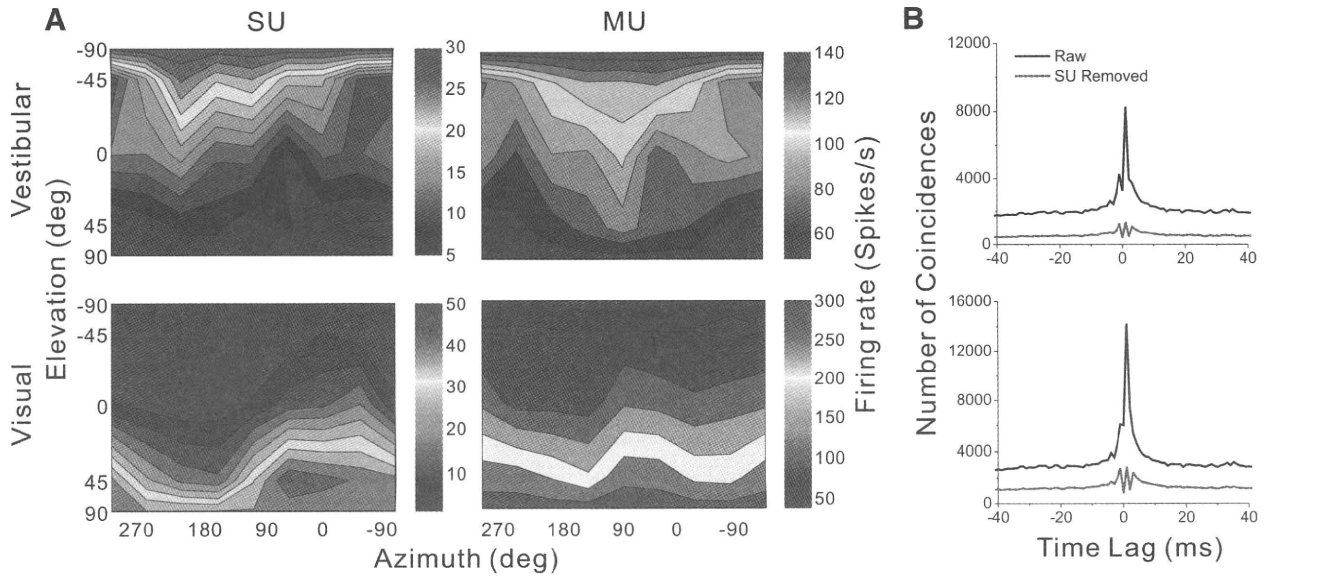

FIG. 3. Example data from an MSTd neuron tested in the vestibular (*top row*) and visual (*bottom row*) rotation conditions. Format as described for Fig. 2.

Figure 4 summarizes the results of the cross-correlation analyses between SU and MU responses before and after removing SU spikes from the MU signal. Because we found no significant differences in the cross-correlation results between translation and rotation protocols, we computed average cross-correlograms across all data sets. As shown in Fig. 4, there was a sharp correlation peak at a lag of 0 ms before SUs were removed and the amplitude of this peak greatly exceeded the SD of the baseline correlation (at large time lags). After SUs were removed, the correlogram was relatively flat and the residual peak was within 1SD of the baseline correlation. Therefore our pruning procedure (see METHODS) was largely effective in removing the artificial coupling between SU and MU signals—we feel confident that any similarity in tuning between SU and MU responses reflects neural clustering in MSTd. Note, however, that there may also be genuine neural coupling between SU and MU signals that takes the form of a broader, shallower peak around *time 0*. Since we removed events from the MU signal only within ±1 ms of SU spikes, some of these broad correlation peaks remain.

We next summarize the similarity of SU and MU tuning across the population of MSTd neurons. To summarize response strength, we computed the difference between the peak response and spontaneous activity (R_{max} − spont). Figure 5*A* shows this metric for SU and MU responses recorded simultaneously during the translation protocol. The vast majority of data points lie above the diagonal for both the vestibular (red) and visual (cyan) conditions. The average MU/SU peak response ratios were 4.1 and 3.4 for the vestibular and visual conditions, respectively. Figure 5*B* shows analogous data for the rotation protocol. The average MU/SU peak response ratios were 3.6 and 3.1 for the vestibular and visual rotation conditions, respectively. Thus for both the translation and rotation protocols, the average MU response was significantly stronger than the average SU response (paired *t*-test, $P \ll 0.001$). Thus even if we failed to exclude every SU spike from the MU activity, any residual SU activity would account for little of the MU response.

To quantify the strength of 3D spatial tuning, we computed a direction discrimination index (DDI; see METHODS) that quantified the peak to trough modulation of the neuron's tuning

TABLE 2. *Percentage of SU and MU with significant spatial tuning for rotation (ANOVA, $P \leq 0.05$)*

	MU	
SU	$P \leq 0.05$	$P > 0.05$
Vestibular (n = 81)		
$P \leq 0.05$	56/81 (69%)	19/81 (23%)
$P > 0.05$	1/81 (1%)	5/81 (6%)
Visual (n = 66)		
$P \leq 0.05$	57/66 (86%)	8/66 (12%)
$P > 0.05$	1/66 (2%)	0/66 (0%)

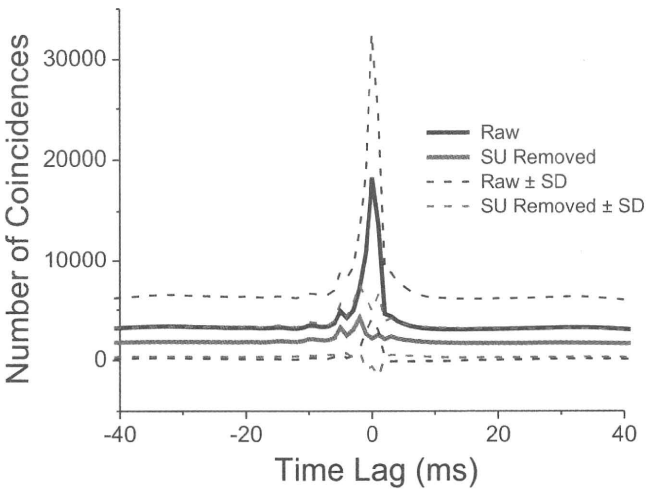

FIG. 4. Averaged cross-correlation function for all pairs of SU and MU tested using the translation (n = 287) or rotation (n = 81) protocols. Translation and rotation data have been pooled. The solid black line shows the raw average cross-correlation before SU spikes were excluded from the MU signal. The dashed black lines represent ±1 SD around this mean. The solid red line shows the average cross-correlation function after SU spikes were removed from the MU signal. Dashed red lines represent ±1 SD around this mean.

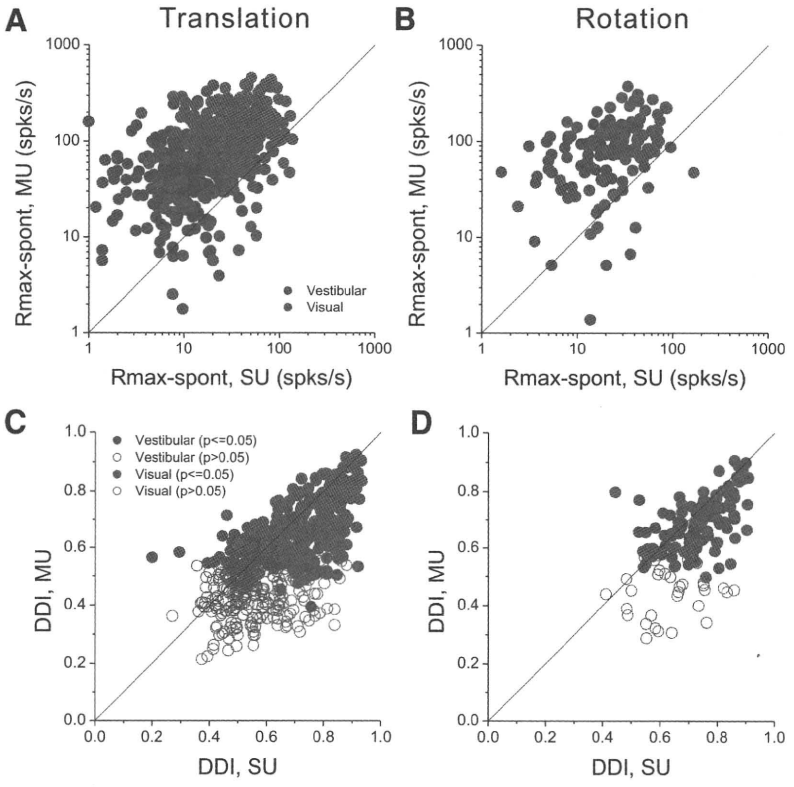

FIG. 5. Quantitative summary of peak response (R_{max} − spont) and tuning strength (direction discrimination index [DDI]) derived from MU and SU responses. In each panel, data from SUs are plotted on the x-axis and data from the corresponding MU responses are plotted on the y-axis. Red and cyan symbols represent data from vestibular and visual conditions, respectively. *A*: comparison of SU and MU peak responses for the vestibular translation (n = 287) and visual translation (n = 272) conditions. *B*: comparison of SU and MU peak responses for the vestibular rotation (n = 81) and visual rotation (n = 66) conditions. *C*: comparison of SU and MU DDI values for the translation conditions. Closed (open) circles: neurons with (without) significant MU tuning (ANOVA, $P < 0.05$). *D*: comparison of SU and MU DDIs for the rotation conditions.

profile relative to the noise level (see also Takahashi et al. 2007). Figure 5*C* compares DDI values from SU and MU responses measured during the translation protocol (red: vestibular; cyan: visual). Filled symbols represent data sets with significant MU tuning (ANOVA, $P < 0.05$), whereas open symbols indicate nonsignificant MU tuning (ANOVA, $P > 0.05$). Most data points tend to fall near or slightly below the unity-slope diagonal. As a result, the mean DDI values for SU activity in the vestibular (0.54) and visual (0.76) conditions were significantly larger than the corresponding mean DDI values for MU activity in the vestibular (0.48) and visual (0.66) conditions (Wilcoxon rank-sum tests, $P < 0.001$). MU and SU DDI values were significantly correlated for both the vestibular ($r = 0.32$; $P < 0.0001$) and visual ($r = 0.51$; $P < 0.0001$) translation conditions, indicating that SUs tend to exhibit weaker tuning at sites where the MU response is poorly tuned. These findings argue against the possibility that flat MU tuning results solely from a combination of SUs that are individually well tuned, but to different directions. Rather, it appears that weak MU tuning is associated, to some degree, with weaker SU tuning.

Figure 5*D* shows a similar comparison of SU and MU DDI values for the rotation protocol. Again, mean DDI values for SU activity significantly exceeded those for MU activity in both the vestibular (0.66 vs. 0.58) and visual (0.77 vs. 0.69) conditions (Wilcoxon rank-sum tests, $P < 0.001$). Despite the smaller sample sizes, MU and SU DDI values were again significantly correlated for both vestibular ($r = 0.48$; $P < 0.0001$) and visual ($r = 0.41$; $P < 0.001$) conditions. Overall, the extent of clustering appears to be comparable for rotation and translation.

We next consider the matching of 3D direction preferences between SU and MU activity. For the translation protocol, 72 data sets (25%) had significant tuning for both SU and MU responses in the vestibular condition, whereas 219 data sets (81%) showed significant SU and MU tuning in the visual condition. For these data sets, we computed the smallest angle in 3D between the preferred direction vectors for SU and MU activity ($|\Delta$ preferred direction$|$). If SU and MU responses have similar direction preferences, this metric should tend toward zero; if there is no clustering of SU and MU preferences, the distribution of $|\Delta$ preferred direction$|$ should be uniform as plotted here. Figure 6*A* shows the distribution of $|\Delta$ preferred direction$|$ for the vestibular (red) and visual (cyan) translation conditions. Both distributions were significantly nonuniform (permutation test, $P < 0.001$), with peaks close to 0°. For the rotation protocol, analogous data were available from 56 recordings in the vestibular condition and 57 recordings in the visual condition. As shown in Fig. 6*B*, the distributions of $|\Delta$ preferred direction$|$ were again clearly nonuniform (permutation test, $P < 0.001$), with peaks close to 0°. Thus when both SU and MU activity in MSTd are significantly tuned, the preferred direction vectors tend to be very similar. Occasionally, there are large differences in direction preference between SU and MU responses and these might occur, for example, when the electrode is located near the boundary between two clusters of neurons that have nearly opposite preferences. Such reversals in direction preference have been observed in area MT, for example (Albright et al. 1984; DeAngelis and Newsome 1999; Malonek et al. 1994).

To further quantify the overall similarity of SU and MU tuning, we computed correlation coefficients between SU and MU response profiles (e.g., Fig. 2*A*). Distributions of the correlation coefficient, R_{SU-MU}, are shown in Fig. 6*C* for the translation conditions. For visual translation, most values of R_{SU-MU} (228/270, 84%) are significantly different from zero

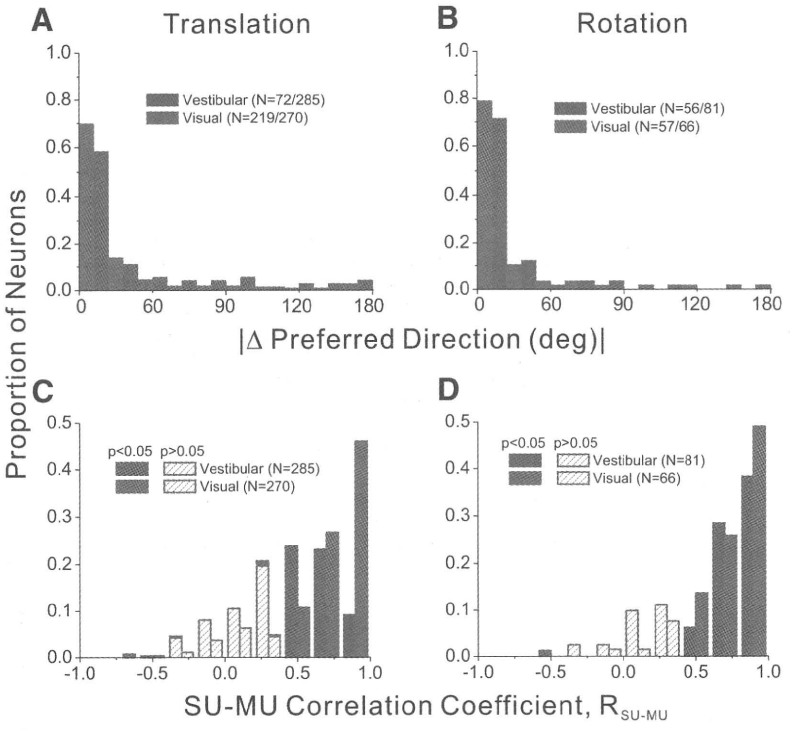

FIG. 6. Summary of differences in direction tuning between SU and MU responses. In *A* and *B*, histograms show the distribution of the difference in 3D preferred directions, | Δ preferred direction |, between corresponding SU and MU responses (| Δ preferred direction | is computed as the smallest angle between the pair of preferred direction vectors in 3D). *A*: histograms of | Δ preferred direction | between SU and MU for the vestibular translation condition (n = 72/285, magenta) and the visual translation condition (n = 219/270, cyan). Note that data are shown here only for recordings in which both SU and MU responses showed significant tuning. *B*: histograms of | Δ preferred direction | for the vestibular rotation (n = 56/81) and visual rotation (n = 57/66) conditions. *C* and *D* show distributions of correlation coefficients (R_{SU-MU}) computed by comparing SU and MU tuning profiles. *C*: distributions of R_{SU-MU} are shown for the vestibular translation (n = 285) and visual translation (n = 270) conditions. Filled bars denote values of R_{SU-MU} that are significantly different from zero. *D*: distributions of R_{SU-MU} are shown for the vestibular rotation (n = 81) and visual rotation (n = 66) conditions.

(filled cyan bars), with an overall median value of 0.76. For vestibular translation, about half of the correlation coefficients (164/285, 57%) are significantly different from zero (filled red bars), with an overall median value of 0.46. This correlation analysis further supports the finding that clustering of selectivity is stronger for visual translation than vestibular translation. Figure 6*D* shows analogous data for the rotation conditions. For visual rotation, 58/66 data sets (89%) have R_{SU-MU} values significantly different from zero, with an overall median value of 0.80. For vestibular rotation, the values are similar: 60/81 R_{SU-MU} values are significantly different from zero (74%), with an overall median value of 0.73.

The analyses of Fig. 6 consider the similarity of direction tuning between SU and MU responses separately for the visual and vestibular conditions. We now consider the congruency of visual and vestibular selectivity. Figure 7*A* shows the distribution of differences in direction preference between visual and vestibular responses for SUs recorded in the translation condition. As reported previously (Fetsch et al. 2007; Gu et al. 2006, 2007), this distribution is clearly bimodal. Roughly half of MSTd neurons have similar visual and vestibular translation preferences ("congruent" cells) and roughly half have widely different preferences ("opposite" cells). A very similar pattern is seen for MU activity from the translation condition, as shown in Fig. 7*B*, indicating that these two extremes of visual/vestibular congruency are preserved in MU activity. For the subset of recordings with significant translation tuning in both visual and vestibular responses, Fig. 7*C* compares the congruency of SU and MU tuning. With the exception of a handful of recording sites, the congruency of MU responses generally matches well with the congruency of SU responses (R = 0.57, $P < 0.001$). Thus both congruent and opposite cells appear to be clustered in area MSTd.

Figure 7, *D–F* shows the analogous congruency data for the rotation experiment. As reported previously (Takahashi et al. 2007), almost all SUs are opposite cells in response to rotation stimuli (Fig. 7*D*) and MU responses in area MSTd exhibit a very similar pattern of results (Fig. 7*E*). The SU and MU measures of congruency again generally agree well (Fig. 7*F*; R = 0.34, P = 0.024), even though there is relatively little variation in these data across recording sites compared with the translation condition.

To provide a simple graphical summary of the results quantified in Figs. 5–7, we computed population tuning curves for SU and MU activity. To simplify the presentation, we constructed tuning curves from the subset of stimulus directions that lie in the horizontal plane (azimuth varies with elevation fixed at zero). This yields clear tuning for most neurons (Fetsch et al. 2007; Gu et al. 2006) and allows us to average tuning curves across neurons. Open symbols in Fig. 8*A* show the average tuning curve for SUs in the vestibular translation condition. For each SU with significant tuning (ANOVA, $P <$ 0.05), the data were shifted such that the maximal response occurs at zero azimuth, spontaneous activity was subtracted, and the resulting curves were averaged across neurons. For comparison, filled symbols in Fig. 8*A* show the average tuning curve for the corresponding MU responses. In this case, spontaneous activity was again subtracted from each MU curve and the MU data were aligned to the azimuth preferred by each SU. Thus if there was no clustering of tuning in MSTd, the MU population tuning curve should be flat. As seen in Fig. 8*A*, average MU responses to vestibular translation have tuning consistent with SU responses, although the modulation depth of this MU tuning is modest and the error bars are substantially larger than those of the SU activity. By comparison, in the visual translation condition (Fig. 8*B*), average MU responses have considerably stronger modulation consistent with the analyses of Fig. 5. Finally, population tuning curves in Fig. 8, *C* and *D* show average selectivity for SU and MU responses in the vestibular and visual rotation conditions, respectively. In

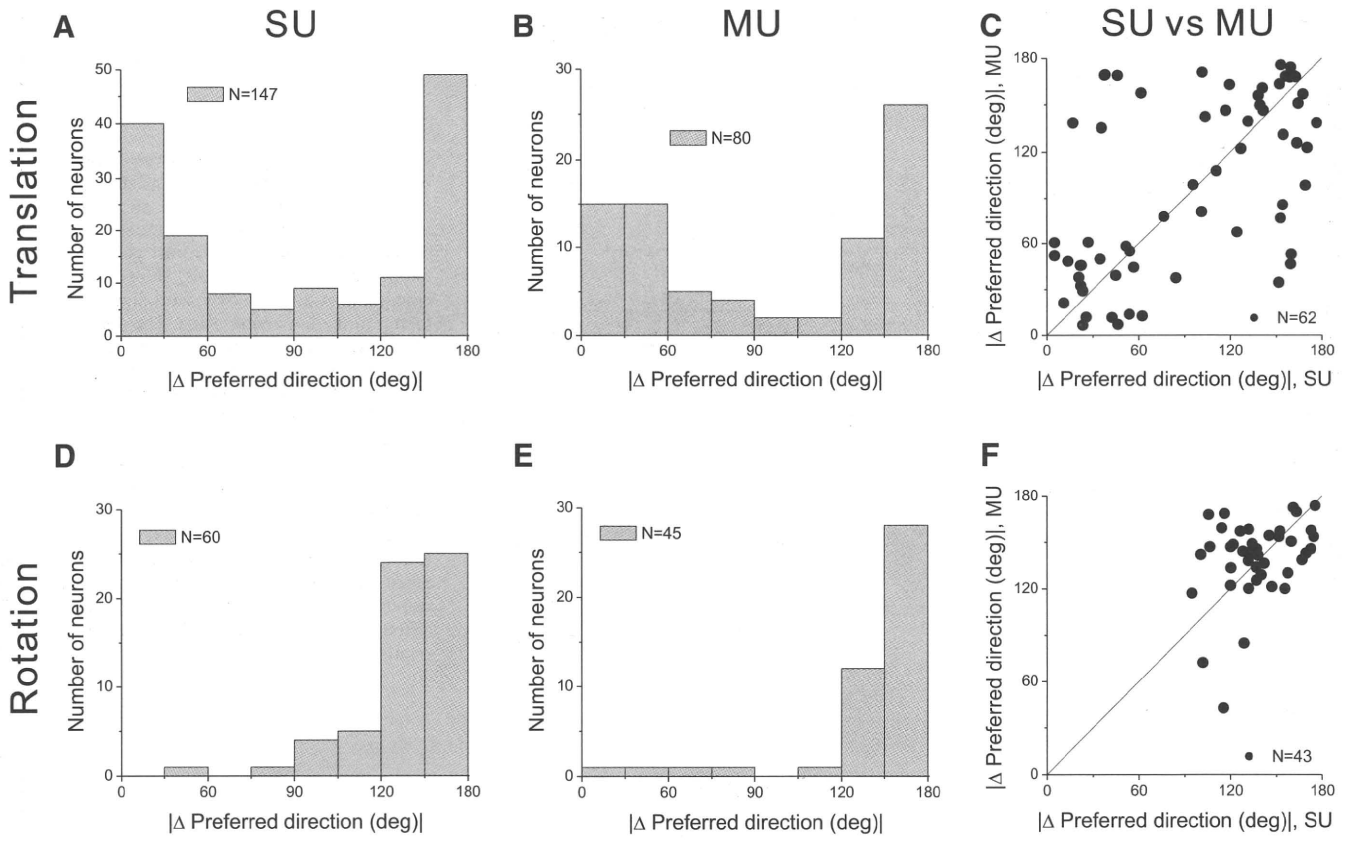

FIG. 7. Congruency of visual and vestibular tuning in SU and MU responses. *A*: congruency of SU responses for translation. The distribution of differences in 3D direction preference between visual and vestibular responses (| Δ preferred direction |) is bimodal, indicating roughly equal numbers of "congruent" and "opposite" cells. Data are shown for the subset of SUs with significant direction tuning in both visual and vestibular conditions. *B*: analogous data are shown for MU activity measured during the translation protocol. *C*: congruency (| Δ preferred direction |) for MU responses is plotted against that for SU responses. Data are shown for recording sites with significant SU and MU tuning for both visual and vestibular conditions. *D*: congruency of SU responses for the rotation protocol. *E*: congruency of MU responses to rotation stimuli. *F*: comparison of congruency between SU and MU responses obtained during the rotation protocol.

both rotation conditions, average MU responses show strong tuning that aligns well with the SU responses.

Clustering of receptive field properties in MSTd

The measurements described earlier strongly suggest that neurons in area MSTd are clustered according to their 3D preferences for translational and rotational movements. Because the optic flow stimuli were large-field (90 × 90°) and the 3D tuning may reflect a complex interaction of visual motion signals across space, we were also interested in whether more basic visual response properties (receptive fields) are clustered in MSTd. Thus for a subset of 70 recordings, we used a reverse-correlation technique (see METHODS) to quantitatively characterize the directional RF structure of SU and MU responses.

Figure 9*A* shows the RF map obtained for an example SU. The visual stimulus was a dynamic random-dot display consisting of a 4 × 4 grid of subfields that spanned the 90 × 90° visual display (Fig. 1*C*). Each subfield contained a random-dot pattern moving coherently in one of eight possible directions (0°: rightward; 90°: upward). Every 100 ms, the direction of motion was chosen randomly for each of the 16 subfields, such that each subregion of the receptive field was probed with all eight directions of motion. The resulting spike train was cross-correlated to the stimulus sequence for each subfield, resulting in a distinct direction–time map for each individual subfield (Fig. 9*A*). If the portion of the neuron's RF overlying a particular stimulus subfield produces a direction-selective visual response, then that subfield will show structure in the direction–time map. If no directional response is elicited by the stimulus in a particular subfield, the direction–time map will be unstructured. For the example neuron in Fig. 9*A*, the central four subfields show clear directional structure, with a preference for rightward and slightly upward motion. Note that this MSTd RF is clearly bilateral and includes the fovea. Figure 9*B* shows the RF map for simultaneously recorded MU activity. The pattern of selectivity for the MU is nearly identical, suggesting that RF properties are locally clustered.

To further quantify the similarity of RF structure between SU and MU responses, we reduced each direction–time map to a single direction tuning curve by taking a cross section through the data at a correlation delay of $T_{peak} \approx 80$ ms (horizontal lines in Fig. 9). This peak delay (T_{peak}) was chosen as the delay that maximized the variance in direction tuning across all subfields of the RF map (estimated separately for SU and MU activities). The resulting direction tuning curves for each subfield are shown in Fig. 10 for the same example recording (MU: filled symbols; SU: open symbols). Solid and dashed curves show the best fits of a wrapped Gaussian function (see METHODS). Two characteristics were used to quantify each direction tuning

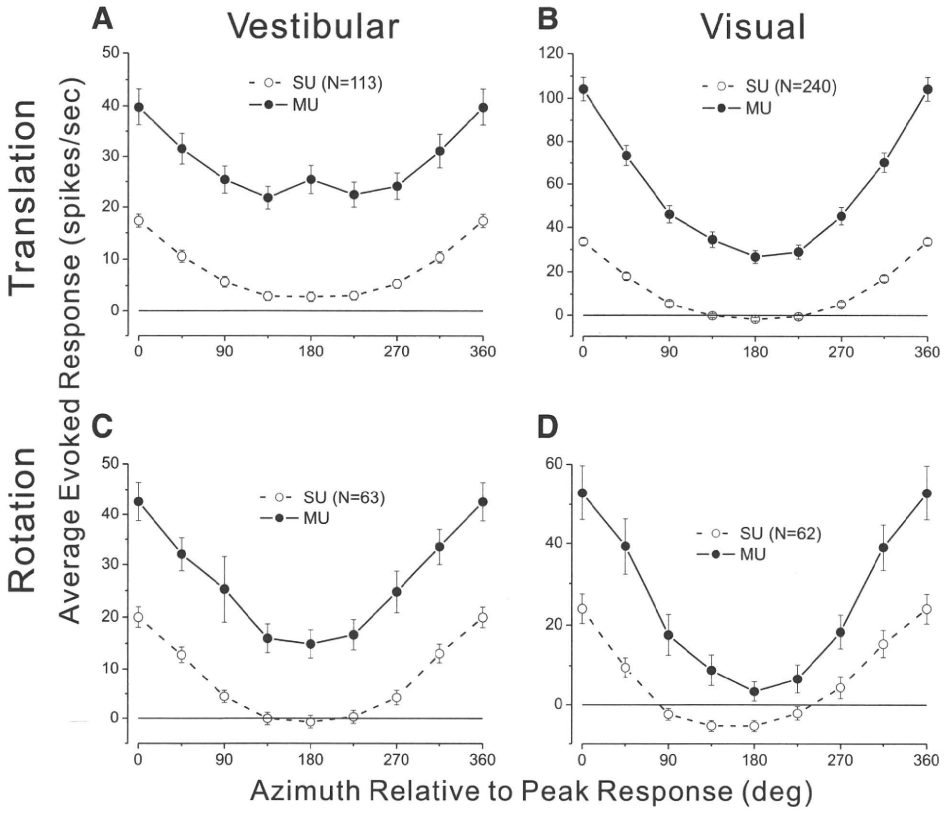

FIG. 8. Summary tuning curves comparing SU and MU responses to translation and rotation stimuli. For each neuron, an azimuth tuning curve was obtained from the subset of directions lying in the horizontal plane. SU tuning curves (open symbols) show the average response of all SUs with significant tuning in the horizontal plane. The data are aligned to the peak response of each SU and spontaneous activity is subtracted before averaging. MU tuning curves (filled symbols) show the average MU responses (after subtracting spontaneous activity). MU data were aligned to the peak response of each corresponding SU. *A*: data from the vestibular translation condition. *B*: data from the visual translation condition. Note that MU tuning is stronger and more reliable compared with vestibular translation. *C*: average responses from the vestibular rotation condition. *D*: average responses from the visual rotation condition.

curve: *1*) the vector sum of normalized responses, which reflects the strength of direction selectivity, and *2*) the preferred direction determined by the peak of the Gaussian fit. For the example data set in Fig. 10, the normalized vector sum is significantly greater than zero for both SU and MU responses in the central four subfields ($P < 0.05$; see METHODS).

Figure 11*A* summarizes the strength of direction tuning (normalized vector sum) for SU and MU responses obtained from each subfield of all 70 data sets ($n = 1{,}120$ total). Filled symbols denote data with significant MU direction tuning ($P < 0.05$). Most of the data points fall below the diagonal, such that the average vector sum for MU activity (0.24) is significantly lower than the average for SU activity (0.33) (paired *t*-test, $P < 0.05$). The vector sum metrics for MU and SU are strongly correlated ($r = 0.71$; $P < 0.001$), indicating that subfields with strong directional selectivity in MU activity also have strong SU tuning.

To summarize the relationship between direction preferences for SU and MU activity, Fig. 11*B* shows the distribution of the difference in preferred directions between SU and MU, as determined from the wrapped Gaussian fits. Data are shown for 351 subfields in which both MU and SU responses had significant direction selectivity. This distribution is clearly nonuniform (permutation test, $P \ll 0.001$), with a peak close to zero (median value: 21.9°). Notably, 61% of SU/MU pairs have direction preferences within 30° of each other, indicating that local visual direction preferences are strongly clustered in MSTd. There is also a hint of a second peak in the distribution near direction differences of 180°, suggesting that SUs occasionally have direction preferences opposite to the MU preference.

Whereas the preceding analysis focused on the direction tuning of SU/MU activity in each subfield of the mapping grid, the reverse-correlation data can also be used to estimate the overall spatial profile of the receptive field. For this purpose, we plotted the amplitude of the wrapped Gaussian fit as a function of the location of each subfield (e.g., Fig. 12*A*). Note that this provides a spatial map of the regions of the RF that produce direction-selective responses; nondirectional portions of the receptive field, should they exist, would not be represented. Figure 12*B* shows the analogous spatial map for the simultaneously recorded MU activity. The SU and MU responses have very similar spatial profiles. To quantify the location and extent of these spatial RFs, the data were fit with a two-dimensional (2D) Gaussian (see METHODS). For the example SU, the best-fitting Gaussian (Fig. 12*C*) is centered at (−1.5°, −5.4°) and has dimensions of 38 × 42° (full width at half-maximum [FWHM], horizontal × vertical). For the corresponding MU profile, the center is located at (2.5°, −0.9°) and the dimensions are 41 × 38°.

Among the 70 data sets obtained using reverse correlation, spatial RF maps with statistically significant structure (permutation test, $P < 0.05$; see METHODS) were obtained for 64% (45/70) of SUs and 66% (46/70) of MUs. Forty-four percent (31/70) had significant spatial structure for both SU and MU responses. Three SU/MU pairs were excluded because of poor fits with the 2D Gaussian function (correlation coefficient, $r < 0.8$, between measured spatial profile and 2D Gaussian fit). For the 28 data sets that met the above-cited criteria, we quantified the degree of overlap of the SU and MU RFs by taking the distance between the centers of the Gaussian fits and normalizing by the average SD of the Gaussian (see METHODS). The distribution of this normalized RF separation, along both horizontal and vertical dimensions, is shown in Fig. 13*A*. If the SU and MU RFs were completely overlapping, these values would all be very close to zero. Mean

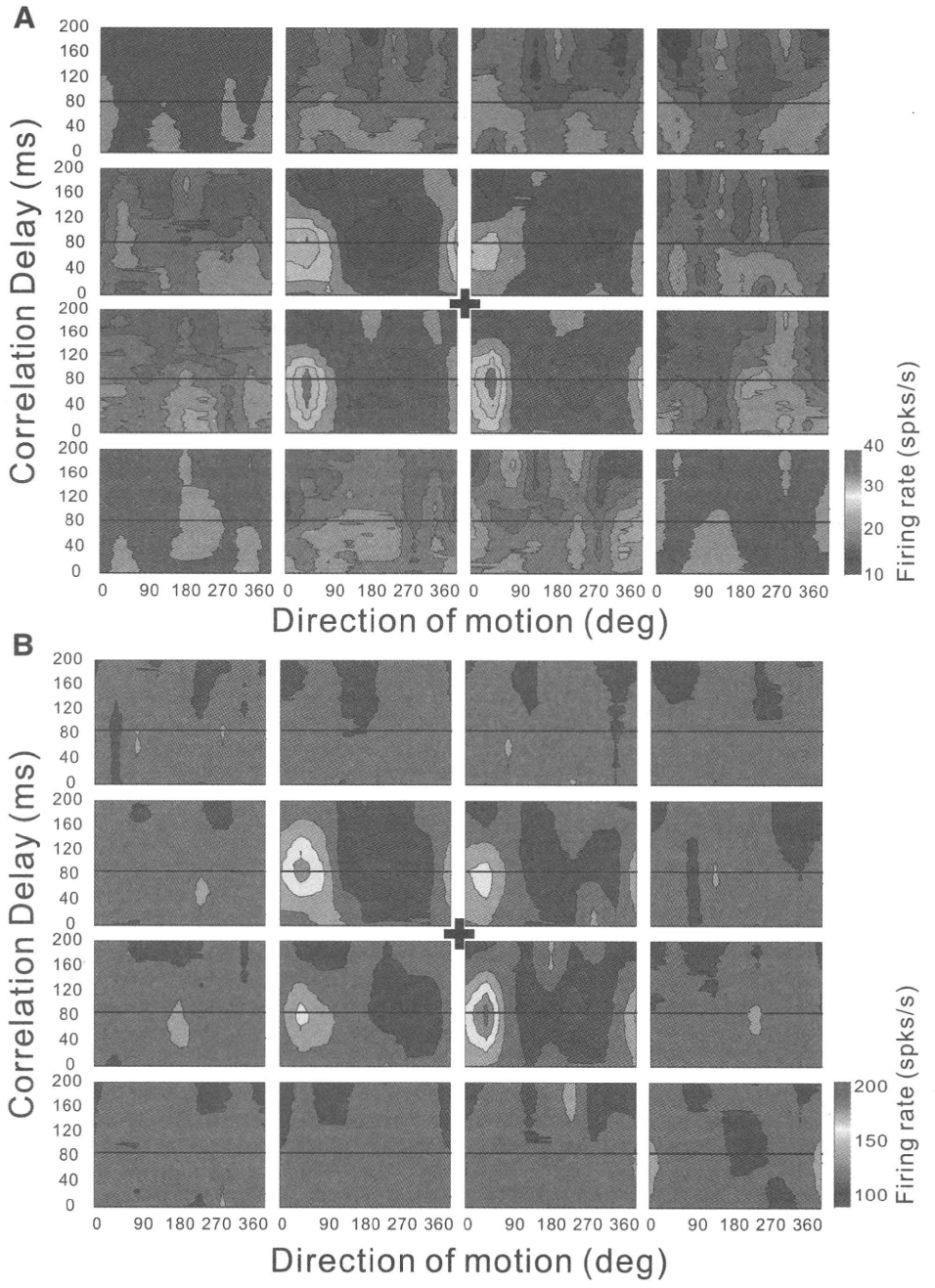

FIG. 9. Direction–time receptive fields (RFs) are shown as contour plots for an example of SU (*A*) and MU (*B*) activity recorded simultaneously. In this case, the stimulated region of the visual field (90 × 90°) was divided into a 4 × 4 grid of 16 locations. Each subfield spanned 15 × 15 cm on the 60 × 60 cm tangent screen (viewed from 30 cm). As a result, the exact angular subtense of each subfield varied somewhat with its location. At each location, the stimulus could move in one of 8 directions. For each location, the neural response is cross-correlated with the motion impulse sequence (see METHODS). The direction of motion is indicated along the horizontal axis (with 0° indicating rightward and 90° indicating upward) and the reverse-correlation delay is indicated along the vertical axis. A "+" in the center denotes the center of the display screen where the monkey fixated. The optimal reverse-correlation delay is indicated by the black horizontal line in each panel.

values are 0.26 ± 0.04 SE for the horizontal separation and 0.27 ± 0.04 SE for vertical. Thus on average the distance between centers of the SU and MU RFs is approximately one fourth of the SD of the Gaussian fit, or about one tenth of the FWHM. This indicates that the spatial profiles of SU and MU responses overlap extensively.

If the RFs of MSTd neurons extend past the boundaries of the 90 × 90° display screen, SU and MU RFs might appear to overlap more extensively than they actually do (thus leading to artificially low values of normalized RF separation). Based on 2D Gaussian fits, we found that 37/70 SUs and 38/70 MUs had RFs that were well contained within the display screen. We computed normalized RF separations for SU/MU pairs with RFs contained within the display screen (n = 25) and compared these to SU/MU pairs in which at least one RF extended off the display screen (n = 8). We found no significant differences between these two groups (t-test, P = 0.29) for horizontal separation. For vertical separation, SU/MU pairs with RFs that extended off the display screen had vertical separations that were significantly greater than SU/MU pairs with RFs contained within the display (P = 0.02). Thus our measures of RF separation were not artificially reduced by RFs extending off the display screen.

Figure 13*B* compares the size of corresponding SU and MU RFs (FWHM along horizontal and vertical dimensions). There was no significant difference between SU and MU responses in terms of these size metrics (paired t-test, $P > 0.05$, n = 28). Together with the separation data of Fig. 13*A*, these results

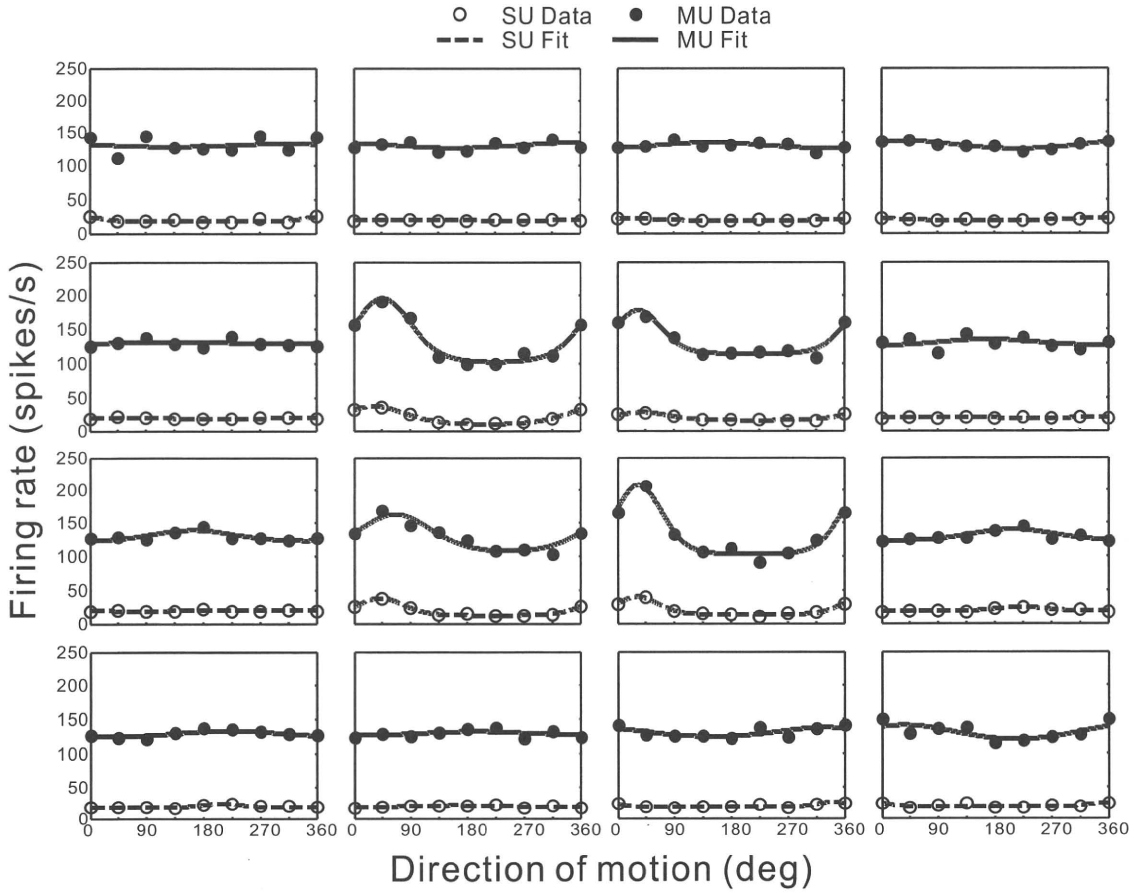

FIG. 10. Direction tuning curves at the peak response latency (T_{peak} = 82 ms) are shown for the same SU and MU data illustrated in Fig. 9. For each location in the stimulus grid, filled (open) circles show the MU (SU) response to 8 different directions of motion, 45° apart. The solid (dashed) curve is the best-fitting wrapped Gaussian function for MU (SU).

indicate that the spatial RFs of SU and MU responses overlap heavily, although it should be noted that our measurements may miss regions of the RFs that are not directionally selective. The average RF size in our population was 39 × 42° for SUs and 36 × 38° for MUs (FWHM dimensions). These values are largely in agreement with previous reports (Desimone and Ungerleider 1986; Komatsu and Wurtz 1988; Van Essen et al. 1981).

DISCUSSION

By comparing the responses of isolated single units with the undifferentiated activity of neighboring neurons, we have examined the clustering of response properties in area MSTd. Our results confirm previous findings of clustered optic flow selectivity in area MSTd (Britten 1998; Duffy and Wurtz 1991; Lagae et al. 1994; Saito et al. 1986) and also show that basic visual receptive field properties are clustered. Our main new finding is that the vestibular sensitivity of MSTd neurons to both translational and rotational motion is also clustered in MSTd. In general, the selectivity of MU activity for vestibular translation was weak compared with visual translation and rotation, as discussed further in the following text. However, when MU signals showed clear selectivity for direction of translation, the tuning was generally consistent with that of simultaneously recorded single units. These findings suggest that there is a topographic organization of vestibular signals in area MSTd.

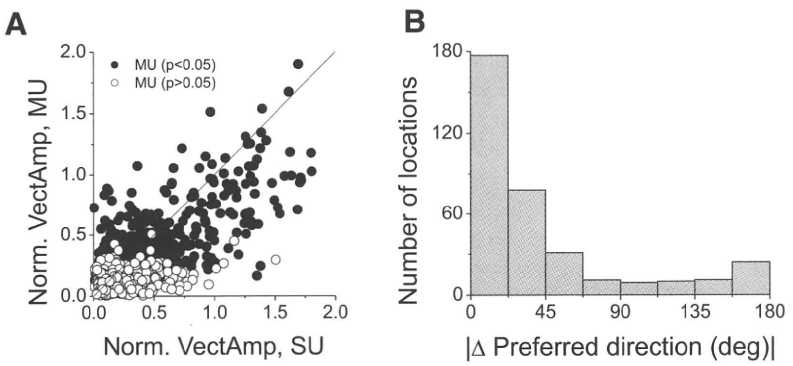

FIG. 11. Summary of differences in direction tuning between SU and MU responses for each location tested within the receptive fields of a population of neurons. *A*: the scatterplot shows the relationship between the vector sum of normalized responses for SU and MU activity. Closed (open) circles: locations with (without) significant direction tuning for MU responses ($P < 0.05$). *B*: histogram of the differences between SU and MU direction preferences for all RF locations at which both signals showed significant tuning.

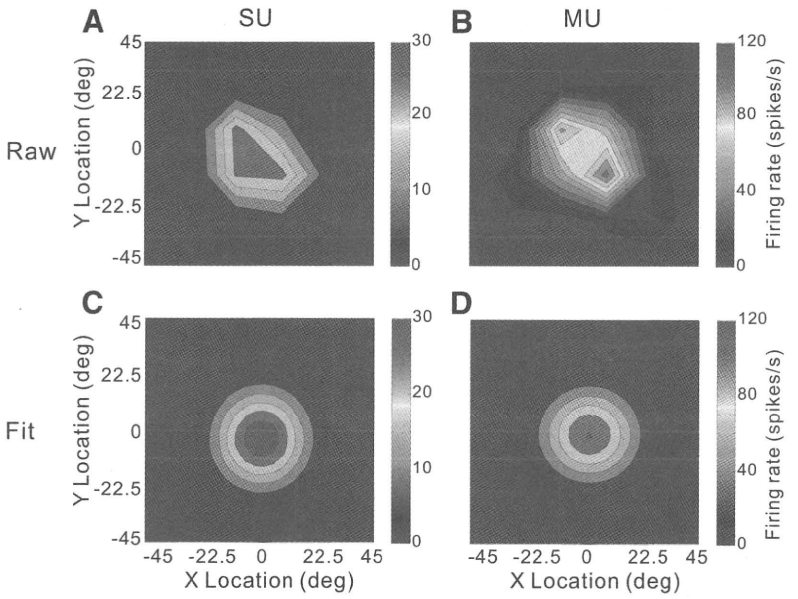

FIG. 12. Example of spatial RF structure derived from reverse-correlation measurements. Two-dimensional (2D) spatial (X–Y) RFs are shown as contour plots for an example data set. Positive x values indicate the right visual hemifield and positive y values indicate the upper visual field. *A*: SU receptive field profile. Each point in the RF map represents the amplitude of the direction tuning curve obtained at the peak response latency. *B*: MU receptive field profile. *C* and *D*: 2D Gaussian fits to the spatial RF profiles of the SU and MU responses, respectively. The full width at half-maximum along the x dimension (FWHMx) is 38° for the SU and 41° for the MU; the full width at half-maximum along the y dimension (FWHMy) is 42° for the SU and 38° for the MU. The center of the RF is located at (−1.5°, −5.4°) for the SU response and (2.5°, −0.89°) for the MU response.

It should be noted, however, that our data limited us to a comparison of SU and MU activity obtained from the same recording site. Thus although our results establish clustering of vestibular selectivity in MSTd, we cannot conclude that MSTd contains a columnar organization of vestibular signals. Although clustering is usually associated with columns (Mountcastle 1957, 1997), it may also be possible to have clustering without columns (Liu and Newsome 2003). We also cannot directly infer anything about the topographic organization of vestibular signals across the surface of the cortex because we did not record neural activity at multiple locations along electrode penetrations through MSTd (see Britten 1998). Our results therefore leave open a number of aspects of the organization of vestibular signals in MSTd, but we do clearly establish clustering of vestibular signals and compare this to clustering of optic flow responses.

The remainder of the discussion will consider *1*) the strength of visual clustering in our data with respect to previous findings and *2*) the difference in strength of clustering between visual and vestibular responses, particularly for translational motion.

Clustering of optic flow selectivity in area MSTd

In his quantitative study of clustering of optic flow properties in MSTd, Britten (1998) measured MU or SU tuning at intervals of 50–100 μm along electrode penetrations through MSTd. He correlated the optic flow tuning observed at a particular recording site with the tuning observed at other sites along the penetration and this revealed a modest, but statistically significant, tendency for optic flow tuning to be clustered. For recording sites separated by <500 μm the optic flow tuning curves were significantly correlated for about 30% of cases. Beyond 500 μm, significant correlations between tuning curves were rare.

The heading data of Britten (1998) (his Fig. 3*A*) can best be compared with our visual translation condition. In our data, we found significant MU tuning for >80% of recording sites and the heading preferences of SU and MU responses were closely matched in the vast majority of cases (Fig. 6*A*). Thus it may appear that we have found stronger evidence for clustering of optic flow selectivity than did Britten (1998). However, a couple of methodological differences may account for most of this apparent difference. First, all of our neurons were tested with a range of headings that sampled all possible directions in 3D (Fig. 1*A*), whereas Britten measured heading tuning curves over a much more restricted range of headings (around straightforward) in the horizontal plane (typically −40 to +40°). For recording sites tested over this restricted range, a weak correlation between tuning curves may result from either site having weak selectivity over the range of headings tested. Second, intersite distance may be a factor. Britten always

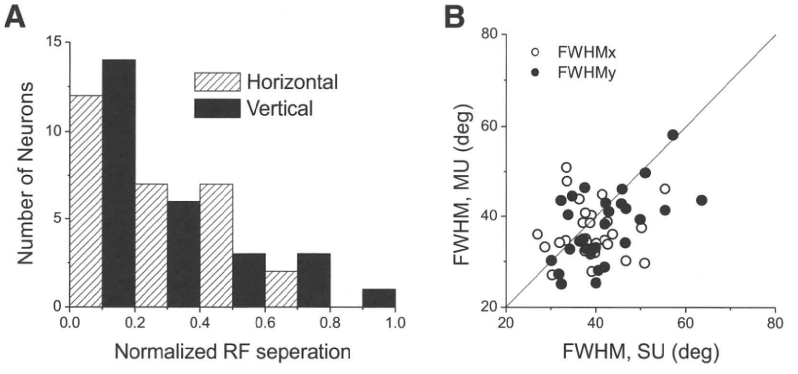

FIG. 13. Summary of RF overlap between SU and MU responses. *A*: distribution of the normalized RF separation, which is defined as the distance between the centers of the SU and MU RFs normalized by the average width of the SU and MU RFs. Hatched and filled bars show normalized separations measured along the horizontal and vertical dimensions, respectively. *B*: comparison of receptive field sizes, FWHMx (open symbols) and FWHMy (filled symbols), between SU and MU responses.

compared heading tuning at sites separated by some distance, whereas we compared SU and MU activity from the same location. Given that correlations between tuning curves fall off with distance, as shown by Britten, it is not surprising that the strongest similarities in tuning would be observed among neurons close to the tip of the electrode. Given these methodological differences, our results do not appear to be inconsistent with Britten's.

Differences in clustering between visual and vestibular responses

When MU responses show significant tuning, there is generally good agreement between SU and MU preferences for both visual and vestibular stimuli (Fig. 6). This appears to suggest that visual and vestibular responses are clustered to a similar degree. For vestibular translation, on the other hand, only 32% of MU recordings show significant heading tuning, whereas 82% of MU recordings show significant tuning for the visual translation condition. If SU and MU preferences for vestibular translation are generally well matched (Fig. 6*A*), then why do MU responses seldom show significant tuning? One possible explanation is simply that SU responses to vestibular translation are less selective than SU responses to visual translation, as can be seen in Fig. 5*C*. Indeed, only 56% of SUs show significant tuning for vestibular translation, whereas 97% of SUs are significantly tuned for visual translation. However, when a SU is significantly tuned, the MU activity is also tuned in only 45% of cases for vestibular translation versus 84% for visual translation.

Another possibility is that weaker selectivity in MU responses to vestibular translation may be affected by the visual–vestibular congruency of tuning for SUs. As shown previously and in Fig. 7, *A* and *B*, approximately half of MSTd neurons have congruent preferences for visual and vestibular translation, whereas the other half have opposite preferences (Gu et al. 2006, 2007). This creates an interesting issue for functional organization of these neurons because opposite cells could be clustered with congruent cells according to either their visual preference or their vestibular preference (this issue is much less relevant for rotation because there are very few congruent cells for rotation; Takahashi et al. 2007). If, for example, MSTd neurons were predominantly clustered according to their visual preference for translation, then the vestibular preferences of nearby congruent and opposite neurons will not be consistent. This would contribute to weak MU tuning in the vestibular translation condition while preserving strong MU tuning in the visual translation condition. At present, our data do not allow us to resolve this issue. Additional experiments, involving recording from multiple isolated SUs, may be necessary to fully understand how congruent and opposite neurons are organized within MSTd. This issue may have important consequences for understanding how the responses of congruent and opposite neurons are decoded during perceptual tasks.

ACKNOWLEDGMENTS

We thank A. Turner and E. White for excellent monkey care and training.

GRANTS

This work was supported by National Eye Institute Grants EY-017866 to D. E. Angelaki and EY-016178 and an EJLB Foundation (Canada) grant to G. C. DeAngelis.

REFERENCES

Albright TD, Desimone R, Gross CG. Columnar organization of directionally selective cells in visual area MT of the macaque. *J Neurophysiol* 51: 16–31, 1984.

Borghuis BG, Perge JA, Vajda I, van Wezel RJ, van de Grind WA, Lankheet MJ. The motion reverse correlation (MRC) method: a linear systems approach in the motion domain. *J Neurosci Methods* 123: 153–166, 2003.

Born RT, Bradley DC. Structure and function of visual area MT. *Annu Rev Neurosci* 28: 157–189, 2005.

Born RT, Tootell RB. Segregation of global and local motion processing in primate middle temporal visual area. *Nature* 357: 497–499, 1992.

Bremmer F, Kubischik M, Pekel M, Lappe M, Hoffmann KP. Linear vestibular self-motion signals in monkey medial superior temporal area. *Ann NY Acad Sci* 871: 272–281, 1999.

Britten KH. Clustering of response selectivity in the medial superior temporal area of extrastriate cortex in the macaque monkey. *Vis Neurosci* 15: 553–558, 1998.

Churchland AK, Lisberger SG. Relationship between extraretinal component of firing rate and eye speed in area MST of macaque monkeys. *J Neurophysiol* 94: 2416–2426, 2005.

DeAngelis GC, Newsome WT. Organization of disparity-selective neurons in macaque area MT. *J Neurosci* 19: 1398–1415, 1999.

DeAngelis GC, Ohzawa I, Freeman RD. Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *J Neurophysiol* 69: 1091–1117, 1993.

DeAngelis GC, Uka T. Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89: 1094–1111, 2003.

Desimone R, Ungerleider LG. Multiple visual areas in the caudal superior temporal sulcus of the macaque. *J Comp Neurol* 248: 164–189, 1986.

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

Duffy CJ, Wurtz RH. Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *J Neurophysiol* 65: 1329–1345, 1991.

Duffy CJ, Wurtz RH. Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci* 15: 5192–5208, 1995.

Eckhorn R, Krause F, Nelson JI. The RF-cinematogram. A cross-correlation technique for mapping several visual receptive fields at once. *Biol Cybern* 69: 37–55, 1993.

Fetsch CR, Wang S, Gu Y, Deangelis GC, Angelaki DE. Spatial reference frames of visual, vestibular, and multimodal heading signals in the dorsal subdivision of the medial superior temporal area. *J Neurosci* 27: 700–712, 2007.

Froehler MT, Duffy CJ. Cortical neurons encoding path and place: where you go is where you are. *Science* 295: 2462–2465, 2002.

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

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

Horton JC, Adams DL. The cortical column: a structure without a function. *Philos Trans R Soc Lond B Biol Sci* 360: 837–862, 2005.

Jones JP, Palmer LA. The two-dimensional spatial structure of simple receptive fields in cat striate cortex. *J Neurophysiol* 58: 1187–1211, 1987.

Komatsu H, Wurtz RH. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J Neurophysiol* 60: 580–603, 1988.

Lagae L, Maes H, Raiguel S, Xiao DK, Orban GA. Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. *J Neurophysiol* 71: 1597–1626, 1994.

Liu J, Newsome WT. Functional organization of speed tuned neurons in visual area MT. *J Neurophysiol* 89: 246–256, 2003.

Malonek D, Tootell RB, Grinvald A. Optical imaging reveals the functional architecture of neurons processing shape and motion in owl monkey area MT. *Proc Biol Sci* 258: 109–119, 1994.

Mountcastle V. Modality and topographic properties of single neurons of cat's somatic sensory area. *J Neurophysiol* 20: 408–438, 1957.

Mountcastle VB. The columnar organization of the neocortex. *Brain* 120: 701–722, 1997.

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

Prince SJ, Pointon AD, Cumming BG, Parker AJ. Quantitative analysis of the responses of V1 neurons to horizontal disparity in dynamic random-dot stereograms. *J Neurophysiol* 87: 191–208, 2002.

Ringach DL, Sapiro G, Shapley R. A subspace reverse-correlation technique for the study of visual neurons. *Vision Res* 37: 2455–2464, 1997.

Saito H, Yukie M, Tanaka K, Hikosaka K, Fukada Y, Iwai E. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J Neurosci* 6: 145–157, 1986.

Snyder JP. *Map Projections: A Working Manual*. Washington, DC: U.S. Government Printing Office, 1987, p. 182–190.

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

Van Essen DC, Maunsell JH, Bixby JL. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J Comp Neurol* 199: 293–326, 1981.

Wennekers T. Orientation tuning properties of simple cells in area V1 derived from an approximate analysis of nonlinear neural field models. *Neural Comput* 13: 1721–1747, 2001.

Yang T, Maunsell JH. The effect of perceptual learning on neuronal responses in monkey visual area V4. *J Neurosci* 24: 1617–1626, 2004.

Zeki SM. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol* 236: 549–573, 1974.

Behavioral/Systems/Cognitive

Does the Middle Temporal Area Carry Vestibular Signals Related to Self-Motion?

Syed A. Chowdhury,[1] Katsumasa Takahashi,[1] Gregory C. DeAngelis,[2]* and Dora E. Angelaki[1]*

[1]Department of 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, New York 14627

Recent studies have described vestibular responses in the dorsal medial superior temporal area (MSTd), a region of extrastriate visual cortex thought to be involved in self-motion perception. The pathways by which vestibular signals are conveyed to area MSTd are currently unclear, and one possibility is that vestibular signals are already present in areas that are known to provide visual inputs to MSTd. Thus, we examined whether selective vestibular responses are exhibited by single neurons in the middle temporal area (MT), a visual motion-sensitive region that projects heavily to area MSTd. We compared responses in MT and MSTd to three-dimensional rotational and translational stimuli that were either presented using a motion platform (vestibular condition) or simulated using optic flow (visual condition). When monkeys fixated a visual target generated by a projector, half of MT cells (and most MSTd neurons) showed significant tuning during the vestibular rotation condition. However, when the fixation target was generated by a laser in a dark room, most MT neurons lost their vestibular tuning whereas most MSTd neurons retained their selectivity. Similar results were obtained for free viewing in darkness. Our findings indicate that MT neurons do not show genuine vestibular responses to self-motion; rather, their tuning in the vestibular rotation condition can be explained by retinal slip due to a residual vestibulo-ocular reflex. Thus, the robust vestibular signals observed in area MSTd do not arise through inputs from area MT.

Introduction

Self-motion perception relies heavily on both visual motion (optic flow) and vestibular cues (for review, see Angelaki et al., 2009). Neurons with directional selectivity for both visual and vestibular stimuli are strong candidates to mediate multisensory integration for self-motion perception. In particular, neurons in both the dorsal portion of the medial superior temporal area (MSTd) and the ventral intraparietal area (VIP) show selectivity to complex optic flow patterns experienced during self-motion (Tanaka et al., 1986; Duffy and Wurtz, 1991, 1995; Graziano et al., 1994; Schaafsma and Duysens, 1996; Bremmer et al., 2002a), and also respond selectively to translation and rotation of the body in darkness (Duffy, 1998; Bremmer et al., 1999, 2002b; Froehler and Duffy, 2002; Schlack et al., 2002; Gu et al., 2006; Takahashi et al., 2007). Recent experiments have linked area MSTd to perception of heading based on vestibular signals (Gu et al., 2007), and have implicated MSTd neurons in combining optic flow and vestibular cues to enhance heading discrimination (Gu et al., 2008).

Despite considerable focus on the roles of MSTd and VIP in visual-vestibular integration, the pathways by which vestibular signals reach these areas remain unknown. One hypothesis is that visual and vestibular inputs arrive at MSTd/VIP through distinct pathways, including a vestibular pathway that remains unidentified. An alternative hypothesis is that vestibular signals are already present in areas that are known to provide visual inputs to MSTd and VIP. The middle temporal (MT) area is an important visual motion processing region in the superior temporal sulcus (STS) (Born and Bradley, 2005), and projects heavily to other functional areas in macaques, including MSTd and VIP (Maunsell and van Essen, 1983; Ungerleider and Desimone, 1986). Area MT has mainly been studied in relation to purely visual functions (Born and Bradley, 2005), and MT neurons are generally not thought to receive strong extra-retinal inputs. For example, MT neurons were reported to show little response during smooth pursuit eye movements, whereas MST neurons were often strongly modulated by pursuit (Komatsu and Wurtz, 1988; Newsome et al., 1988). However, a recent study showed that responses of many MT neurons to visual stimuli were modulated by extra-retinal signals, which serve to disambiguate the depth sign of motion parallax (Nadler et al., 2008). This finding could reflect either vestibular or eye movement inputs to MT. Hence, we investigated whether MT neurons respond to translational and rotational vestibular stimulation.

We used an experimental protocol similar to that used previously to characterize visual and vestibular selectivity in MSTd (Gu et al., 2006; Takahashi et al., 2007). In control experiments, we measured vestibular tuning in complete darkness or during fixation of a laser-generated fixation point in a dark room. We find that MT responses are not modulated by vestibular stimulation in the absence of retinal image motion. We conclude that, although visual motion signals in MSTd likely arise through pro-

Received Jan. 1, 2009; revised March 18, 2009; accepted Aug. 16, 2009.

This work was supported by National Institutes of Health (NIH) Grants EY017866 and EY019087 (to D.E.A.) and NIH Grant EY016178 (to G.C.D.). We thank Amanda Turner and Erin White for excellent monkey care and training, Babatunde Adeyemo for eye movement analyses, and Yong Gu for valuable advice throughout these experiments.

*G.C.D. and D.E.A. contributed equally to this work.

Correspondence should be addressed to Dr. Gregory C. DeAngelis, Center for Visual Science, University of Rochester, 245 Meliora Hall, Rochester, NY 14627. E-mail: gdeangelis@cvs.rochester.edu.

DOI:10.1523/JNEUROSCI.0004-09.2009

 0270-6474/09/2912020-11$15.00/0

jections from MT (Maunsell and van Essen, 1983; Ungerleider and Desimone, 1986), vestibular information reaches MSTd through distinct, as yet uncharacterized, routes.

Materials and Methods

Responses of MT and MSTd neurons were recorded from three adult rhesus monkeys (*Macaca Mulatta*) chronically implanted with a head restraint ring and a scleral search coil to monitor eye movements. A recording grid was placed inside the ring and used to guide electrode penetrations into the superior temporal sulcus (for details, see Gu et al., 2006). All animal procedures were approved by the Washington University Animal Care and Use Committee and fully conformed to National Institutes of Heath guidelines. Monkeys were trained using operant conditioning to fixate visual targets for fluid rewards.

Single neurons were isolated using a conventional amplifier, a band-pass eight-pole filter (400–5000 Hz), and a dual voltage–time window discriminator (BAK Electronics). The times of occurrence of action potentials and all behavioral events were recorded with 1 ms resolution by the data acquisition computer. Eye movement traces were low-pass filtered and sampled at 250 Hz. Raw neural signals were also digitized at 25 kHz and stored to disk for off-line spike sorting and additional analyses.

Areas MT and MSTd were identified based on the characteristic patterns of gray and white matter transitions along electrode penetrations, as examined with respect to magnetic resonance imaging scans, and based on the response properties of single neurons and multiunit activity (for details, see Gu et al., 2006). Receptive fields of MT/MSTd neurons were mapped 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. In some neurons, quantitative receptive field maps were also obtained (see Fig. 4) using a reverse correlation method as described below. We recorded from any MT/MSTd neuron that was spontaneously active or that responded to a large field of flickering random dots.

Vestibular and visual stimuli. Data were collected in a virtual reality apparatus (for details, see Gu et al., 2006; Takahashi et al., 2007) that consists of a six-degree of freedom motion platform (MOOG 6DOF2000E; Moog) and a stereoscopic video projector (Mirage 2000; Christie Digital Systems). Animals viewed visual stimuli projected onto a tangent screen (60 × 60 cm) at a viewing distance of 32 cm, resulting in a field of view of ~90 × 90°. Because the apparatus was enclosed on all sides with black, opaque material, there was no change in the visual image during movement, other than visual stimuli presented on the tangent screen and small retinal slip induced by fixational eye movements. The visual stimulus consisted of a three-dimensional (3D) cloud of "stars," which was generated by an OpenGL graphics board (nVidia Quadro FX3000G) having a pixel resolution of 1280 × 1024.

The "rotation" protocol consisted of real or simulated rotations (right hand rule) around 26 distinct axes separated by 45° in both azimuth and elevation (see Fig. 1, inset). This included all combinations of movement vectors having 8 different azimuth angles (0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°) and 3 different elevation angles: 0° (the horizontal plane), −45°, and 45° (8 × 3 = 24 directions). For example, azimuth angles of 0° and 180° (elevation = 0°) correspond to pitch-up and pitch-down rotations, respectively. Azimuth angles of 90° and 270° (elevation = 0°) correspond to roll rotations (right-ear-down and left-ear-down, respectively). Two additional stimulus conditions had elevation angles of –90° or 90°, corresponding to leftward or rightward yaw rotation, respectively (bringing the total number of directions to 26). Each movement trajectory had a duration of 2 s and consisted of a Gaussian velocity profile. Rotation amplitude was 9° and peak angular velocity was 23.5°/s. Five repetitions of each motion direction were randomly interleaved, along with a null condition in which the motion platform remained stationary and only a fixation point appeared on the display screen (to assess spontaneous activity). Note that the rotation stimuli were generated such that all axes of rotation passed through a common point that was located in the mid-sagittal plane of the head along the interaural axis. Thus, the animal was always rotated around this point at the center of the head.

The "translation" protocol consisted of straight translational movements along the same 26 directions described above for the rotation protocol. Again, each movement trajectory had a duration of 2 s with a Gaussian velocity profile. Translation 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.

The rotation and translation protocols both included two stimulus conditions. (1) In the "vestibular" condition, the monkey was moved in the absence of optic flow. The screen was blank, except for a fixation point that remained at a fixed head-centered location throughout the motion trajectory (i.e., the fixation point moved with the animal). (2) In the "visual" condition, the motion platform was stationary, while optic flow simulating movement through a cloud of stars was presented on the screen. In initial experiments, MT neurons were tested with both vestibular and visual (optic flow) versions of the rotation and translation protocols, interleaved within the same block of trials (for details, see Takahashi et al., 2007).

To further explore the origins of response selectivity observed under the vestibular rotation condition, subsequent experiments concentrated only on the vestibular rotation stimulus, which was delivered under 3 different viewing conditions, in separate blocks of trials (135 trials per block). The first viewing condition, which we refer to as the 'projector' condition, was identical to that used in previous studies to characterize the vestibular responses of MSTd neurons (Gu et al., 2006; Takahashi et al., 2007). In this case, the animal foveated a central, head-fixed target that was projected onto the screen by the Mirage projector. The projected fixation target was a small (0.2° × 0.2°) yellow square with a luminance of 76 cd/m^2. In this condition, the animal was not in complete darkness because of the background illumination of the projector that was 1.8 cd/m^2. The background image produced by the Mirage DLP projector has a faint but visible "screendoor" pattern. Thus, the visual background in the projector condition did contain a subtle visual texture.

In the second viewing condition, which we refer to as the "laser" condition, the projector was turned off and the animal was in darkness, except for a small fixation point that was projected onto the display screen by a head-fixed red laser. The laser-projected fixation target was also ~0.2° × 0.2° in size, and had an apparent luminance similar to that in the projector condition despite the different spectral content. Even following extensive dark adaptation, the display screen was not visible to human observers in the laser condition. Correspondingly, measurements with a Tektronix J17 photometer did not reveal a measurable luminance greater than zero.

For both the projector and laser conditions, the animal was required to foveate the fixation target for 200 ms before the onset of the motion stimulus (fixation windows spanned 2 × 2° of visual angle). The animals were rewarded at the end of each trial for maintaining fixation throughout the stimulus presentation. If fixation was broken at any time during the stimulus, the trial was aborted and the data were discarded. Finally, the third viewing condition, which we refer to as the "darkness" condition, consisted of motion in complete darkness; both the projector and laser were turned off and the animal was freely allowed to make eye movements. In this block of trials, there was no behavioral requirement to fixate, and rewards were delivered manually to keep the animal alert.

We did not collect data using the translation protocol in all of the above viewing conditions because we found little modulation of MT responses during vestibular translation in the projector condition (see Fig. 2*C*,*D*; supplemental Fig. 1, available at www.jneurosci.org as supplemental material).

Data analysis. Analysis and statistical tests were performed in Matlab (MathWorks) using custom scripts. For each stimulus direction and motion type, we computed neural firing rate during the middle l s interval of each trial and averaged across stimulus repetitions to compute the mean firing rate (Gu et al., 2006). Mean responses were then plotted as a function of azimuth and elevation to create 3D tuning functions. To plot these spherical data on cartesian axes, the data were transformed using the Lambert cylindrical equal-area projection (Snyder, 1987), where the abscissa represents azimuth angle and the ordinate corresponds to a sinusoidally transformed version of the elevation angle.

Statistical significance of directional selectivity was assessed using one-way ANOVA for each neuron.

The neuron's preferred direction for each stimulus condition was described by the azimuth and elevation of the vector sum of the individual responses (after subtracting spontaneous activity). In this computation, the mean firing rate in each trial is considered to represent the magnitude of a 3D vector, the direction of which was defined by the azimuth and elevation angles of the particular stimulus (Gu et al., 2006). To quantify the strength of spatial tuning, we computed a direction discrimination index (DDI) (Prince et al., 2002; DeAngelis and Uka, 2003), as follows:

Discrimination index

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

where R_{max} and R_{min} are the mean firing rates of the neuron along the directions that elicited maximal and minimal responses, respectively. SSE is the sum squared error around the mean responses, N is the total number of observations (trials), and M is the number of stimulus directions ($M = 26$). This index quantifies the amount of response modulation (due to changes in stimulus direction) relative to the noise level.

Eye movement analysis. In the projector and laser conditions, animals were required to maintain fixation on a head-fixed target during rotation or translation. Thus, no eye movements were required for the animal to maintain fixation, and the animal should suppress reflexive eye movements driven by the vestibulo-ocular reflex (VOR). Failure to fully suppress the VOR in these stimulus conditions would lead to retinal slip that could elicit visual responses that might be mistakenly interpreted as vestibular responses. Thus, it is important to characterize any residual eye movements that occur in the fixation trials, as well as the more substantial eye movements that would be expected in the darkness condition due to the VOR.

Because torsional eye movements were not measured in these experiments, we could not quantify eye movements that might occur in response to components of roll rotation. Thus, we focused our eye movement analyses on rotations and translations about the 8 stimulus directions within the fronto-parallel plane (yaw and pitch rotations; lateral and vertical translations), as these movements may be expected to elicit horizontal and vertical eye movements. Eye velocity was computed by differentiating filtered eye position traces (boxcar filter, 50 ms width). Fast phases, including saccades, were removed by excluding periods where absolute eye velocity exceeded 6°/s. Data from each individual run were then averaged across stimulus repetitions into mean horizontal and vertical eye velocity profiles (see Fig. 3*A*,*B*).

To quantify the direction and speed of eye movements relative to stimulus motion, we measured the time-average horizontal and vertical components of eye velocity within a 400 ms time window (0.9–1.3 s poststimulus onset) that generally encompassed the peak eye velocity (see Fig. 3*A*). The eye velocity components were then averaged across stimulus repetitions for each distinct direction of motion. Thus, for each recording session, we obtained an average eye velocity vector for each different direction of motion within the fronto-parallel plane. A set of such eye movement vectors for one animal is shown in Figure 3*C* for vestibular rotation in the projector condition.

Receptive field mapping and reverse correlation analysis. It was important to determine the receptive field locations of MT neurons relative to

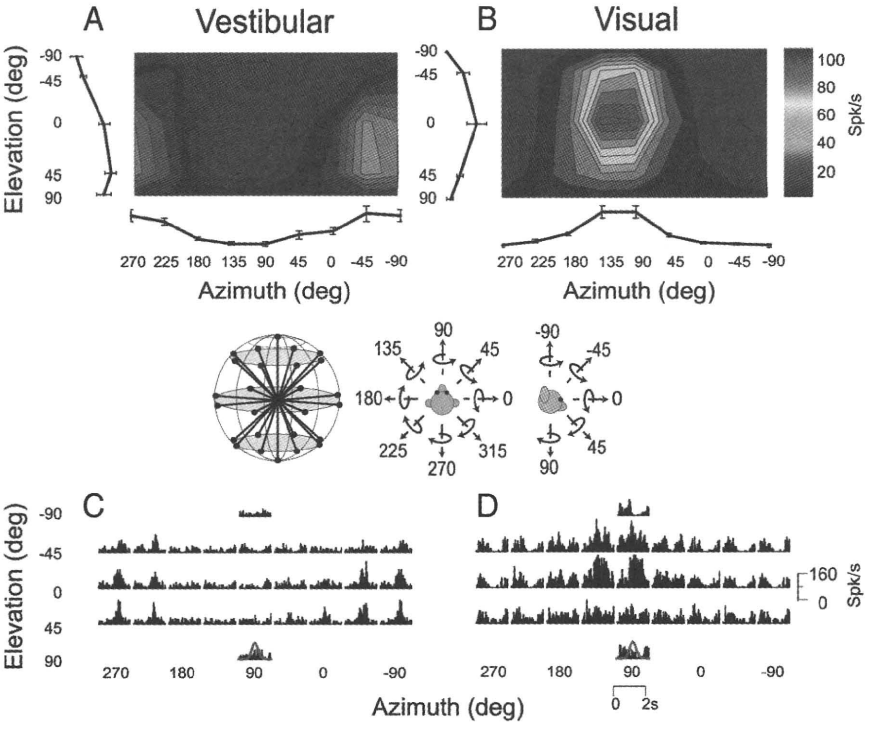

Figure 1. ***A–D***, 3D rotation tuning for an MT neuron tested during the vestibular condition (***A***, ***C***), and the visual condition (***B***, ***D***). The fixation point and visual stimuli were generated by a video projector in this case. Color contour maps in ***A*** and ***B*** show mean firing rate plotted as a function of 26 distinct azimuth and elevation angles (see inset). Each contour map shows the Lambert cylindrical equal-area projection of the original spherical data (Gu et al., 2006). Tuning curves along the margins of each color map illustrate mean ± SEM firing rates plotted as a function of either elevation or azimuth (averaged across azimuth or elevation, respectively). The PSTHs in ***C*** and ***D*** illustrate the corresponding temporal response profiles (each PSTH is 2 s in duration). The red Gaussian curves (bottom) illustrate the stimulus velocity profile.

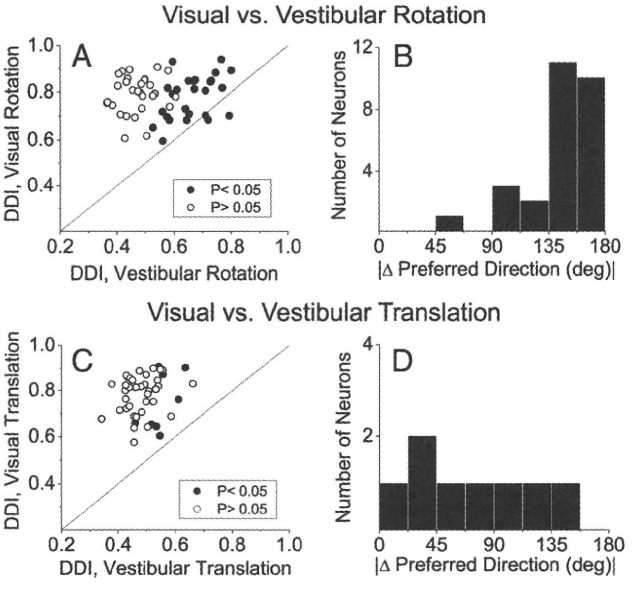

Figure 2. ***A–D***, Summary of the selectivity of MT neurons in response to (***A***, ***B***) visual and vestibular rotation and (***C***, ***D***) visual and vestibular translation. Scatter plots in ***A*** and ***C*** compare the DDI in the visual versus vestibular conditions. Histograms in ***B*** and ***D*** show the absolute difference in 3D direction preferences between visual and vestibular conditions (|Δ preferred direction|) for the rotation and translation protocols, respectively. Data in ***B***, ***D*** are included only for neurons with significant tuning in both stimulus conditions. All data are from the projector condition.

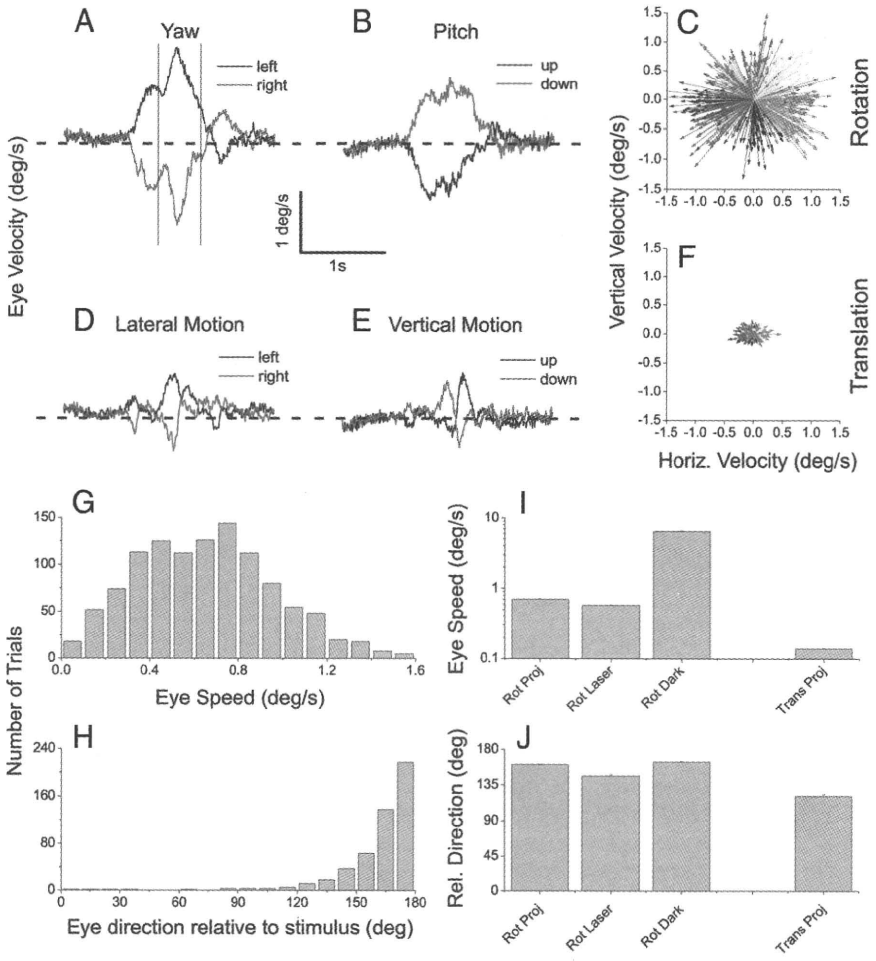

Figure 3. Eye movement analysis showing incomplete suppression of reflexive eye movements. ***A***, Average horizontal eye velocity is shown during leftward (black) and rightward (red) yaw rotation (vestibular condition). Traces show average eye velocity across all of the recording sessions for which data are shown in Figure 2. ***B***, Average vertical eye velocity during upward (black) and downward (red) pitch rotation. ***C***, Vector plot summary of the residual eye velocities in response to vestibular rotation for one monkey in the projector condition. Each vector represents the average eye velocity for one direction of rotation in one experimental session (green, leftward yaw; red, pitch down; purple, rightward yaw; black, pitch up). ***D***, Average horizontal eye velocity during left/right translation. ***E***, Average vertical eye velocity during up/down translation. ***F***, Vector plot summary for the vestibular translation condition for the same animal as in ***C*** (green, left; red, down; purple, right; black, up). ***G***, Distribution of eye speed (vector magnitude) for the same data depicted in ***C***. ***H***, Distribution of eye direction relative to stimulus direction for the data depicted in ***C***. ***I***, Average eye speed across animals for vestibular rotation (projector, laser, and darkness conditions) and vestibular translation (projector condition only). ***J***, Average eye direction relative to stimulus direction for vestibular rotation and translation (format as in ***I***).

both the fixation target and the boundaries of the display screen, as retinal slip of these visual features may be expected to elicit responses. For a subset of MT neurons, we obtained quantitative maps of visual receptive field structure using a multi-patch reverse correlation method that we have previously used to characterize MSTd receptive fields (for details, see Chen et al., 2008).

Briefly, either a 4 × 4 or 6 × 6 grid covered a region of the display screen containing the receptive field of the MT neuron under study. At each location in the grid, a patch of drifting random dots was presented. Dots drifted in one of 8 possible directions, 45 apart, and the direction of motion of each patch was chosen randomly (from a uniform distribution) every 100 ms (6 video frames). The direction of motion of each patch was independent of the other patches presented simultaneously. The speed of motion was generally 40°/s, although this was sometimes reduced for MT neurons that preferred slower speeds.

Responses of MT neurons to this multipatch stimulus were analyzed using reverse correlation, as detailed previously (Chen et al., 2008). For each location in the mapping grid, each spike was assigned to the direction of the stimulus that preceded the spike by T ms, where T is the reverse correlation delay. This was repeated for a range of correlation delays from 0 to 200 ms, yielding a direction-time map for each location in the mapping grid. The correlation delay that produced the maximal variance in the direction maps was then selected, and a direction tuning curve for each grid location was computed at this peak correlation delay. These direction tuning curves were fit with a wrapped Gaussian function, and the amplitude of the tuning curve at each grid location was taken as a measure of neural response for each stimulus location. These values were then plotted as a color contour map (see Fig. 4) to visualize the spatial location and extent of the visual receptive field. These quantitative maps were found to agree well with initial estimates of MT receptive fields based on hand mapping.

Results

Data were collected from 102 MT and 53 MSTd neurons recorded from 6 hemispheres of three monkeys. Two monkeys were used for MT and two for MSTd recordings. One monkey was therefore used for both MT and MSTd recordings. Neurons were not selected according to directional tuning for visual or vestibular stimuli; rather, any neuron was recorded that exhibited spontaneous activity or activity in response to flicker of a large-field random-dot stimulus. In initial recordings from area MT, we used the same experimental protocol used by Takahashi et al. (2007) to study MSTd neurons, hereafter referred to as the projector condition (see Materials and Methods). Figure 1 illustrates data from an example MT neuron tested with the 3D rotation protocol, in which the animal was rotated around 26 axes corresponding to all combinations of azimuth and elevation angles in increments of 45° (see inset). Each movement trajectory, either real (vestibular condition) or visually simulated (visual condition), had a duration of 2 s and consisted of a Gaussian velocity profile. This MT neuron showed clear tuning for both vestibular and visual rotations, as illustrated by the color contour plots of mean firing rate (Fig. 1*A*,*B*) and by the corresponding peristimulus time histograms (PSTHs) (Fig. 1*C*,*D*).

Of 55 MT cells tested under both the vestibular and visual rotation conditions, about half (27, 49%) showed significant tuning in the vestibular condition (ANOVA, $p < 0.05$), whereas all 55 showed significant tuning in the visual condition. Rotational tuning, as assessed by a direction discrimination index (or DDI, see Materials and Methods), was significantly stronger in the visual than the vestibular condition overall (Wilcoxon signed-rank test, $p \ll 0.001$) (Fig. 2*A*). As shown in Figure 2*B*, direction preferences for visual versus vestibular rotation tended to be opposite, as was also seen regularly in area MSTd (Takahashi et al., 2007).

In contrast to the rotation condition, only 8/47 (17%) MT cells showed significant tuning during the vestibular translation

condition (ANOVA, $p < 0.05$). All 47 neurons showed significant visual translation tuning. In general, vestibular translation responses were quite weak, and resulted in DDI values that were much smaller than those elicited by visual translation stimuli (Fig. 2*C*) (Wilcoxon signed-rank test, $p \ll 0.001$). Inspection of PSTHs from all neurons revealed that response modulations were generally not consistent with either stimulus velocity or acceleration. Supplemental Figure 1, available at www.jneurosci.org as supplemental material, shows responses from 4 MT neurons with the largest DDI values in the vestibular translation condition. Among the 8 neurons with weak but significant spatial tuning, the difference in direction preferences between the visual and vestibular translation conditions was relatively uniform and not biased toward 180° (Fig. 2*D*).

The rotational stimuli used here normally generate a robust rotational vestibulo-ocular reflex (RVOR). Although the animals were trained to suppress their RVOR, the 2 × 2° fixation window allows for some residual eye velocity, thus resulting in a visual motion stimulus caused by the fixation point moving relative to the retina. Furthermore, because the DLP projector created substantial background illumination with a very faint screendoor texture (see Materials and Methods), these residual eye movements could activate MT neurons and account for the observed responses in the vestibular rotation condition (Fig. 2*A*, filled symbols).

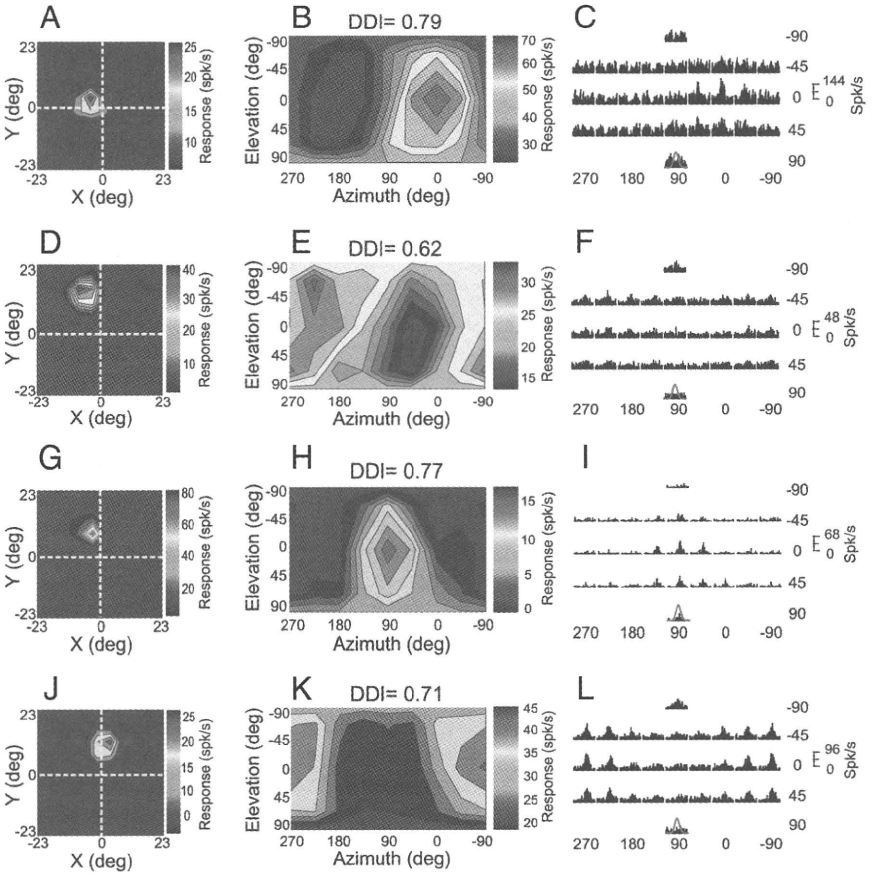

Figure 4. Receptive field maps and vestibular rotation responses for 4 MT neurons tested in the projector condition. ***A***, ***D***, ***G***, ***J***, Receptive field maps as obtained using a multi-patch reverse correlation method (see Materials and Methods). Each color map represents a region of visual space with a length and width one-half as large as the display screen. The fixation point was presented at the intersection of the dashed white lines. The cell depicted in ***A*** had a receptive field that overlapped the fixation point, whereas the other cells did not. ***B***, ***E***, ***H***, ***K***, Vestibular rotation tuning profiles, format as in Figure 1*A*, *B*. ***C***, ***F***, ***I***, ***L***, PSTHs corresponding to the 26 distinct directions of rotation tested in the projector condition. Format as in Figure 1*C*,*D*.

Some aspects of the data in Figure 2 support this interpretation. First, the direction preference of MT cells in the vestibular rotation condition was generally opposite to their visual preference (Fig. 2*B*). Because stimulus directions are referenced to body motion (real or simulated), a vestibular response that is due to retinal slip from an incompletely suppressed RVOR would result in a direction preference opposite to that seen in the visual condition. Second, MT responses were weaker for vestibular translation than vestibular rotation. This result is expected if responses are driven by retinal slip, because the translational VOR (TVOR) is much weaker than the RVOR over the range of stimulus parameters used here (for review, see Angelaki and Hess, 2005). These considerations support the possibility that the observed MT responses to vestibular stimulation might not represent an extra-retinal response to self-motion but rather reflect the exquisite sensitivity of MT neurons to retinal motion slip. Thus, it is critical to characterize the residual eye movements that occur during fixation in the projector condition.

Eye movement analysis

To determine whether there were systematic eye movements in response to vestibular rotation and translation in the projector condition, we analyzed eye traces measured during movements within the fronto-parallel plane (see Materials and Methods for details). Figure 3*A* shows the average horizontal eye velocity across 54 sessions in response to vestibular yaw rotation. Leftward yaw rotation (black trace) results in a rightward eye movement with a peak velocity near 2°/s. Similarly, rightward yaw rotation (red trace) results in leftward eye velocity as expected from an incompletely suppressed RVOR. A similar pattern of eye traces in seen in Figure 3*B* for vestibular pitch rotation. These residual eye movements are summarized for one animal by the vector plot in Figure 3*C*. Each vector represents the average eye velocity for a particular direction of motion in one recording session, measured during a 400 ms time window centered on the peak eye velocity (Fig. 3*A*, vertical lines). For example, the set of green vectors in Figure 3*C* represents eye velocity in response to leftward yaw rotation, and the red vectors represent eye movements in response to pitch down. The distribution of eye speeds (vector magnitudes) corresponding to the data of Figure 3*C* is shown in Figure 3*G*, and it can be seen that the average eye speed is <1°/s. Figure 3*H* shows that the residual eye movements in response to vestibular rotation were generally opposite to the direction of head rotation.

During vestibular translation, some residual eye velocity was also observed (Fig. 3*D*,*E*). Note, however, that these residual eye movements were substantially smaller in amplitude and narrower in time. They were also more biphasic in waveform (Fig. 3*E*) thus resulting in less cumulative retinal slip. This difference is emphasized by the vector plot in Figure 3*F*, which shows that eye

movements for this animal were much smaller in response to vestibular translation and less systematically related to stimulus direction. These data are summarized across animals in Figure 3, *I* and *J*. Mean eye speed in response to translation (0.14°/s) was significantly smaller than mean eye speed (0.71°/s) in response to rotation (Fig. 3*I*, leftmost vs rightmost bars) ($p < 0.001$, *t* test). Moreover, the average directional difference between eye and stimulus was larger (161°) for rotation than translation (121°), as shown in Figure 3*J*.

The eye movement analyses of Figure 3 are consistent with the possibility that significant vestibular rotation tuning in the projector condition may be driven by residual eye movements that are much stronger than those seen in the vestibular translation condition. We next examined whether these vestibular rotation responses depended on receptive field location.

Receptive field analysis

If MT responses in the vestibular rotation condition are driven by retinal slip due to an incompletely suppressed RVOR, then these responses may depend on receptive field location. Retinal slip of the fixation target or the visible screen boundaries might be expected to activate MT neurons with receptive fields that are near the fovea or at large eccentricities, respectively. Alternatively, MT neurons at all eccentricities might respond to retinal slip of the faint visual texture produced by the background luminance of the projector. Our analysis suggests that both of these factors contribute to the observed responses.

Figure 4 shows receptive field maps measured using a reverse correlation technique (left column), vestibular rotation tuning profiles (middle column), and response PSTHs (right column) for 4 example MT neurons. Receptive field maps were obtained using a reverse correlation technique (Chen et al., 2008). The neuron in Figure 4*A–C* has a receptive field that overlaps the fovea (Fig. 4*A*), and shows robust responses in the vestibular rotation condition that are well tuned (Fig. 4*B*,*C*). These responses may be driven by retinal slip of the fixation target. In contrast, the other three neurons depicted in Figure 4 have receptive fields that are well separated from both the fixation target and the boundaries of the projection screen (screen boundaries are at ±45°), yet these neurons still produce robust responses in the vestibular rotation condition. We infer that the responses of these MT neurons are likely driven by retinal slip of the faint background texture produced by the DLP projector.

Figure 5*A* summarizes the relationship between directional selectivity (DDI) in the vestibular rotation condition and receptive field location. The location of data points along the abscissa represents the eccentricity of the center of the receptive field, and the horizontal error bars represent the size of the receptive field (±2 SD of a two-dimensional Gaussian fit). Filled symbols denote MT neurons with significant vestibular rotation tuning ($p < 0.05$), and symbols containing stars indicate the example neurons of Figure 4. There is a significant negative correlation (Spearman $r = -0.50$, $p < 0.01$) between DDI and receptive field eccentricity, indicating that proximity of the receptive field to the fixation point does contribute to vestibular rotation tuning. Nevertheless, there are several neurons (including the 3 examples in Fig. 4*D–L*) that show significant vestibular rotation tuning yet have receptive fields well separated from the fovea and screen boundaries. Thus, retinal slip of both the fixation point and faint visual background both appear to drive MT responses.

If rotation tuning in the projector condition is driven by retinal slip due to an incompletely suppressed RVOR, then the 3D rotation preference of MT neurons should be linked to the 2D visual direction preference of the neurons. For 17 MT neurons with significant vestibular rotation tuning, reverse correlation maps were available to test this prediction. For each of these neurons, we computed the projection of the preferred 3D rotation axis onto the fronto-parallel plane (for 10/17 neurons, the 3D rotation preference was within 40° of the fronto-parallel plane, and none of the neurons had a 3D rotation preference within 30° of the roll axis). By adding 90° to this projected rotation preference (right-hand rule), we predicted the 2D visual motion direction within the fronto-parallel plane that should best activate the neuron, given that residual eye movements are opposite to the fronto-parallel component of rotation (Fig. 3*H*,*J*). This predicted direction preference is plotted on the ordinate in Figure 5*B*. From the reverse correlation map, we extracted the 2D visual direction preference of each neuron by averaging the direction preferences for each location within the map that showed significant directional tuning (Chen et al., 2008). This measured visual direction preference is plotted on the abscissa in Figure 5*B*. Measured and predicted direction preferences were strongly correlated (circular-circular correlation, $r = 0.98$, $p < 0.001$), indicating that the vestibular rotation preference in the projector condition was highly predictable from the

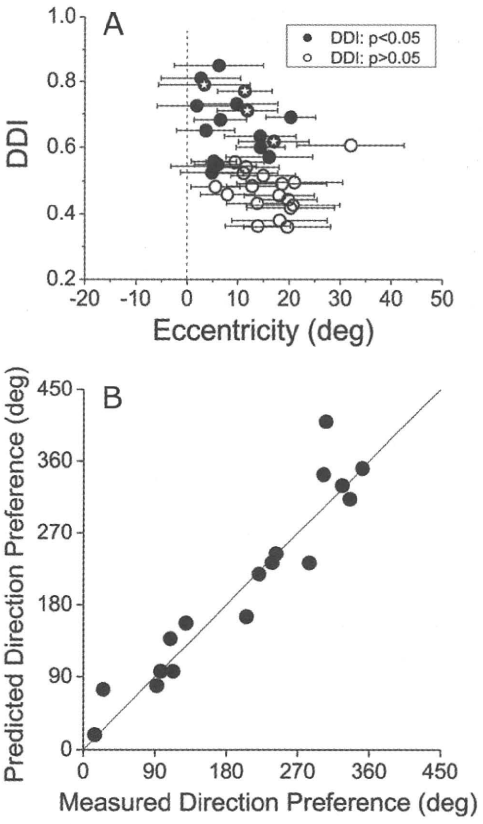

Figure 5. ***A***, Relationship between the strength of directional tuning in the vestibular rotation condition (using the projector) and the eccentricity of MT receptive fields. Each receptive field map, as in Figure 4, was fit with a two-dimensional Gaussian function. The eccentricity of the center of the receptive field is plotted on the abscissa, and the DDI is plotted on the ordinate. Filled symbols denote neurons with statistically significant rotation tuning ($p < 0.05$). Horizontal error bars represent receptive field size as ±2 SD of the Gaussian fit. Thus, the horizontal error bars contain 95% of the area of the receptive field. Symbols filled with stars indicate the 4 example neurons shown in Figure 4. ***B***, Comparison of measured visual direction preferences with predicted preferences from vestibular rotation tuning in the projector condition. See Results for details. The strong correlation suggests that vestibular rotation tuning in the projector condition reflects visual responses to retinal slip.

visual direction preference measured separately using reverse correlation.

Vestibular rotation tuning in MT: projector versus laser and darkness conditions

To further test whether MT responses to vestibular rotation can be explained by retinal slip due to residual RVOR, two additional stimulus conditions were used (see Materials and Methods for details). In the laser condition, the fixation target was generated by a head-fixed laser in an otherwise completely dark room. This condition eliminates the faint visual background texture produced by the projector. In the darkness condition, no fixation point was presented and the animal was in complete darkness. Thus, visual fixation was not required and the animal was allowed to generate an RVOR. The darkness condition should eliminate any remaining responses driven by retinal slip of the fixation point.

Responses of an example MT neuron under both the projector and laser conditions are illustrated in Figure 6. This cell showed clear rotation tuning (Fig. 6*A*) (DDI = 0.61, ANOVA, $p < 0.01$) and stimulus-related PSTHs (Fig. 6*C*) when the fixation target was generated by the projector. When identical rotational motion was delivered while the monkey fixated a laser-generated target in an otherwise dark room, responses were much weaker and rotation tuning was no longer significant (Fig. 6*B*) (DDI = 0.49, ANOVA, $p = 0.16$). The corresponding PSTHs in the laser condition lacked clear stimulus-related temporal modulation (Fig. 6*D*). Given that both the projector and laser conditions involved fixation of a small (0.2 × 0.2°) visual target and that residual eye movements were similar in these two condition (Fig. 3*I*,*J*), these data suggest that most of the response of this neuron to vestibular rotation was driven by retinal slip of the faintly textured background illumination of the projector. This inference is consistent with the observation that this neuron had a receptive field that did not overlap the fixation point or the screen boundaries (Fig. 6*E*).

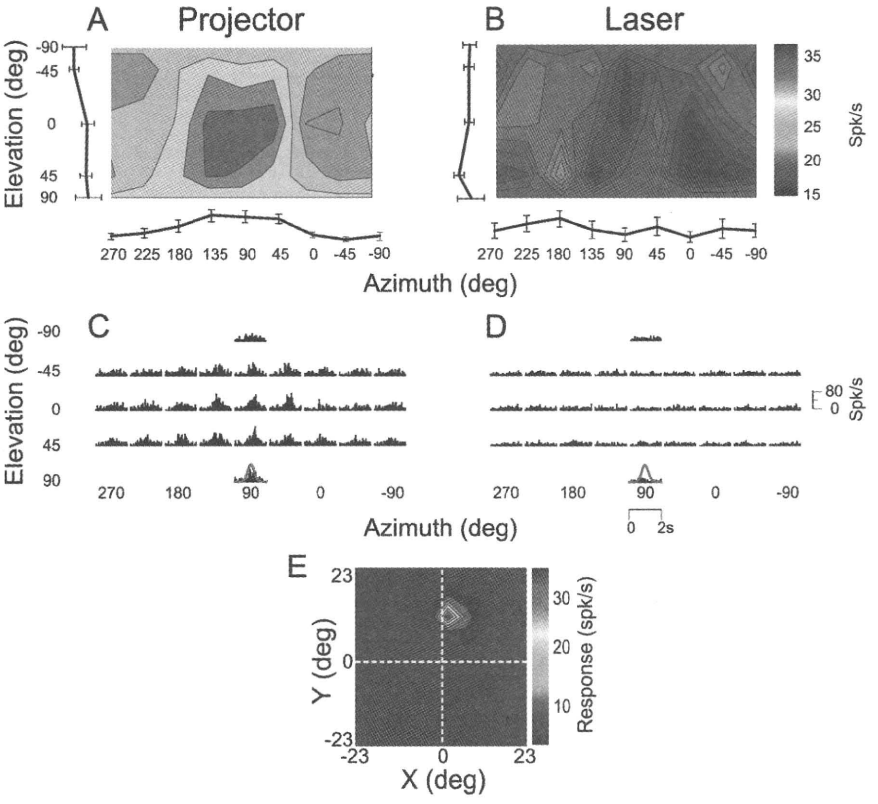

Figure 6. ***A–D***, 3D direction tuning profiles for an MT neuron tested in the vestibular rotation condition when the fixation point is generated by (***A***, ***C***) the video projector or (***B***, ***D***) a head-fixed laser. Color contour maps in ***A*** and ***B*** show the mean firing rate as a function of azimuth and elevation angles (format as in Fig. 1). PSTHs in ***C*** and ***D*** show the corresponding temporal response profiles. ***E***, Receptive field map for this neuron, format as in Figure 4.

We studied the responses of 41 MT cells (15 and 26 from each of the two animals used for MT recordings) under the projector condition and either the laser ($n = 36$) or darkness ($n = 22$) conditions (17 MT cells were tested under all 3 conditions). To quantify the similarity in rotational tuning across conditions, we computed a correlation coefficient between pairs of 3D tuning profiles, such as those shown in Figure 6, *A* and *B*. To obtain this correlation coefficient, the mean response to each direction in the projector condition was plotted against the mean response to the corresponding direction in the laser/darkness condition, and a Pearson correlation coefficient was computed from these data.

Among 36 MT cells tested under both the projector and laser conditions, only 2 (6%) showed a statistically significant correlation ($p < 0.05$) between the two tuning profiles (Fig. 7*A*), and the mean correlation coefficient across neurons was not significantly different from zero (*t* test, $p = 0.29$). This result is further illustrated by plotting the corresponding DDI values in Figure 7*B*. There was no significant correlation between the strength of rotational tuning under the projector and laser conditions, as assessed by DDI ($p = 0.30$). Only 6/36 (16%) MT cells showed significant tuning (ANOVA, $p < 0.05$) in the laser condition, and 4 of these neurons also showed significant tuning in the projector condition (Fig. 7*B*, filled circles). Two of these 4 neurons had receptive fields that overlapped the fovea, and these were the only neurons to show a significant correlation between rotational tuning in the projector and laser conditions (Fig. 7*A*,*B*, red bars/symbols). Responses of these two neurons, which had the two largest DDI values in the laser condition, were thus likely driven by retinal slip of the fixation target. Among the vast majority of neurons (30/36, 84%) with no significant tuning in the laser condition, half showed significant tuning in the projector condition (Fig. 7*B*, filled upright triangles, $n = 15$) and half showed no significant tuning in either the projector or laser conditions (Fig. 7*B*, open circles, $n = 15$).

Figure 7, *C* and *D*, shows the corresponding comparisons between the projector and darkness conditions. Again, the mean correlation coefficient between tuning profiles was not significantly different from zero (*t* test, $p = 0.25$) and only 3/22 MT cells had tuning curves that were significantly correlated between the projector and darkness conditions (Fig. 7*C*, red). In addition, there was no significant correlation between DDI values for the projector and darkness conditions (Fig. 7*D*, $p = 0.60$). Only two MT cells (9%) showed significant rotational tuning in darkness (Fig. 7*D*), and the DDI was rather low for one of these cells (open inverted triangle). The other cell (filled red circle) was an outlier in Figure 7*D*, and we cannot firmly exclude the possibility that this cell was recorded from area MST near the boundary with MT. Among the remaining neurons, 15/22 cells

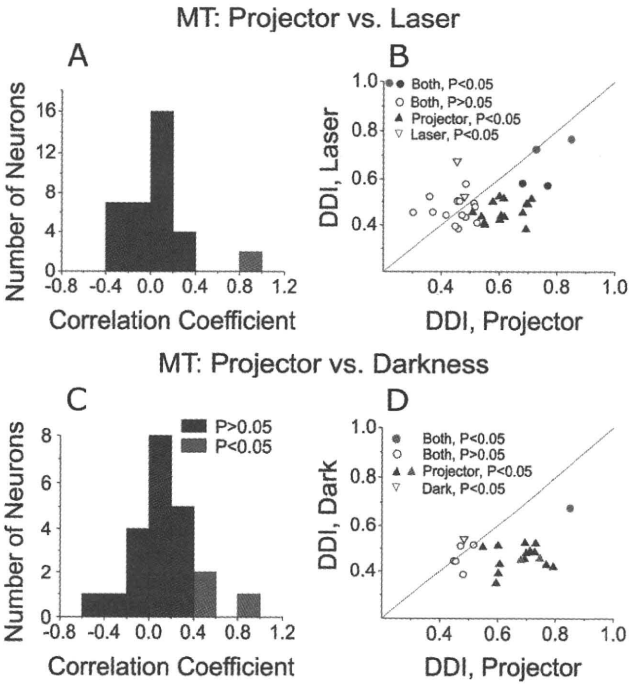

Figure 7. ***A–D***, Comparison of vestibular rotation tuning in MT between projector and laser (***A***, ***B***) or darkness (***C***, ***D***) conditions. ***A***, ***C***, Histograms show the distribution of correlation coefficients between 3D tuning profiles measured in the projector versus laser (***A***) or projector versus darkness (***C***) conditions. Red bars indicate cells with significant ($p < 0.05$) correlations. ***B***, ***D***, Scatter plots compare the DDI between stimulus conditions. Filled circles, Cells with significant tuning (ANOVA, $p < 0.05$) under both conditions. Open circles, Cells with nonsignificant tuning under both conditions. Filled upright triangles, Cells with significant tuning only in the projector condition. Open inverted triangles, Cells with significant tuning only in the laser condition. Red symbols denote cells with significant ($p < 0.05$) correlation coefficients between the projector/laser or projector/darkness conditions.

(68%) showed significant rotation tuning in the projector condition only (Fig. 7*D*, filled triangles) and 5/22 cells showed no significant tuning in either the projector or darkness conditions. Across the population of MT neurons, the median DDI was significantly greater for the projector condition than either the laser or darkness conditions (Wilcoxon matched pairs test, $p < 0.01$ for both comparisons). Overall, MT responses in the vestibular rotation condition were rare when retinal slip of the visual background was removed in the laser and darkness conditions.

Vestibular rotation tuning in MSTd: projector versus laser and darkness conditions

For comparison with MT, we characterized the vestibular rotation tuning of 19 MSTd neurons under both laser and projector conditions, and 48 MSTd neurons under both projector and darkness conditions (14 cells were tested in all 3 conditions). Unlike MT cells, the majority of MSTd neurons showed clear response modulation and consistent 3D rotational tuning under all stimulus conditions. Responses from an example MSTd neuron are shown in Figure 8. This cell showed clear rotation tuning under both the projector and laser conditions, although response strength was a bit lower in the laser condition (Fig. 8, *A* vs *B*; DDI = 0.855 vs 0.731, respectively). Direction preferences were also fairly similar (projector: 89° azimuth and 7° elevation; laser: 123° azimuth and 15° elevation). In addition, response PSTHs largely followed the Gaussian velocity profile of the rotation stimulus under both stimulus conditions (Fig. 8, *C* vs *D*).

Across the population of MSTd neurons studied, rotation tuning profiles measured in the projector condition were often significantly correlated (Pearson correlation, $p < 0.05$) with those measured in the laser condition (Fig. 9*A*, 10/19 significant) or the darkness condition (Fig. 9*C*, 33/48 significant). In both cases, the average correlation coefficient was significantly greater than zero (*t* test, Fig. 9*A*: $p = 0.007$; Fig. 9*C*: $p \ll 0.001$), unlike the pattern of results seen in area MT (Fig. 7*A*,*C*). Similarly, the strength of vestibular rotation tuning in the projector condition (as measured by DDI) was significantly correlated with that measured in the laser (Fig. 9*B*: $p = 0.0016$) and darkness conditions (Fig. 9*D*: $p < 0.001$). When comparing laser and projector conditions, most cells were significantly tuned in both conditions (13/19, 68%; Fig. 9*B*, filled circles) or neither condition (3/19, open circles). Similarly, when comparing projector and darkness conditions, 28 of 48 (58%) MSTd neurons showed significant tuning in both conditions (Fig. 9*D*, filled circles), whereas only 4 cells lacked significant tuning in both conditions. Among the remaining MSTd neurons, 16 of 48 cells (33%) showed a significant DDI in the projector condition but not in darkness. These neurons, like those in MT, may be responding to weak background motion in the projector condition. Thus, the percentage of rotation-selective neurons in the vestibular condition was likely overestimated to some degree by Takahashi et al. (2007).

Discussion

Along with optic flow, vestibular information helps us navigate through space. A number of studies have reported the presence of both optic flow and inertial motion tuning in areas MSTd (Duffy, 1998; Bremmer et al., 1999; Page and Duffy, 2003; Gu et al., 2006; Takahashi et al., 2007) and VIP (Schaafsma and Duysens, 1996; Bremmer et al., 2002a). In particular, two recent studies have emphasized a potential role of MSTd in heading perception based on vestibular signals. First, responses of MSTd neurons show significant trial-by-trial correlations with perceptual decisions about heading that rely strongly on vestibular input (Gu et al., 2007). Second, MSTd responses to inertial motion were abolished after bilateral labyrinthectomy, showing that these responses arise from activation of the vestibular system (Takahashi et al., 2007). In addition, a subpopulation of MSTd neurons with congruent visual and vestibular tuning appears to integrate visual and vestibular cues for improved heading sensitivity (Gu et al., 2008).

Despite the importance of visual/vestibular integration for self-motion perception, the location in the brain where visual and vestibular signals first combine remains unknown. More specifically, the source of the vestibular signals observed in area MSTd is unclear. These signals could arrive in MSTd through the same pathways that carry visual signals or through some separate pathway from other vestibular regions of the brain. Since one of the major projections to areas MSTd and VIP arises from visual area MT (Maunsell and van Essen, 1983; Ungerleider and Desimone, 1986), the present experiments were designed to explore whether MT neurons show directionally selective responses to physical translations and rotations of the subject, as seen previously in area MSTd (Gu et al., 2006, 2007; Takahashi et al., 2007) and area VIP (Bremmer et al., 2002a; Schlack et al., 2002; Klam and Graf, 2006; Chen et al., 2007). Area MT is well known for its important roles in the processing of visual motion (Born and Bradley, 2005), and has been studied extensively with regard to its roles in perception of direction (Britten et al., 1992; Pasternak and Merigan, 1994; Salzman and Newsome, 1994; Purushothaman and Bradley, 2005), perception of speed (Pasternak and Merigan, 1994; Orban

et al., 1995; Priebe and Lisberger, 2004; Liu and Newsome, 2005), perception of depth (DeAngelis et al., 1998; Uka and DeAngelis, 2003, 2004, 2006; Chowdhury and DeAngelis, 2008), and perception of 3D structure from motion and disparity (Xiao et al., 1997; Bradley et al., 1998; Dodd et al., 2001; Vanduffel et al., 2002; Nguyenkim and DeAngelis, 2003). Responses of MT neurons have been generally considered primarily visual in origin. For example, smooth pursuit eye movements were found to robustly modulate responses of MST neurons but not MT neurons (Newsome et al., 1988). However, eye position has been reported to modulate MT responses (Bremmer et al., 1997). In addition, we have recently shown that extraretinal signals modulate responses of MT neurons to code depth sign from motion parallax (Nadler et al., 2008). In that study, both head movements and eye movements were possible sources of extra-retinal inputs. That finding, combined with the close anatomical connectivity between MT and MSTd, further motivated the present study to examine vestibular responses in area MT.

In the same experimental protocol that we used previously to study vestibular rotation responses in MSTd (Takahashi et al., 2007) and VIP (Chen et al., 2007), we found that approximately half of MT neurons were significantly tuned for vestibular rotation. In contrast, only 17% of MT neurons were tuned to vestibular translation in the projector condition, compared with ~60% of MSTd neurons (Gu et al., 2006). As supported by the eye movement data in Figure 3, this difference in strength of vestibular responses to rotation and translation likely arises from differences between the rotational and translational VOR (for review, see Angelaki and Hess, 2005). Whereas the RVOR is robust at low frequencies, the TVOR gain in monkeys is small under the conditions of our experiment: 30 cm viewing distance and a relatively slow motion stimulus (Schwarz and Miles, 1991; Telford et al., 1995, 1997; Angelaki and Hess, 2001; Hess and Angelaki, 2003). Furthermore, because of the unpredictable direction and transient nature of the motion profiles we used, animals could not completely suppress their reflexive eye movements. This incomplete suppression of the VOR results in varying amounts of retinal slip (Fig. 3). Because of the background illumination of the video projector used to generate the head-fixed fixation target (see Materials and Methods), MT neurons would thus be stimulated visually by this retinal slip. This problem would be expected to be larger for rotation, because of the robustness and larger gain of the RVOR under the conditions of our experiment.

To further explore this possibility, we measured vestibular rotation tuning when the head-fixed fixation target was generated by a laser in an otherwise dark room. The results of Figure 7, *A* and *B*, show that much of the tuning seen in the projector condition was due to retinal slip of either the fixation target or the faintly textured background illumination of the video projector. We also recorded MT responses during vestibular rotation in complete darkness, with no fixation requirement. The latter condition eliminates all retinal stimulation, but has the caveat that the eyes are continuously moving during stimulus delivery. Only two MT neurons (9%) showed significant rotation tuning in darkness, just above the number expected by chance. Considering all of the data, we conclude that the rotation selectivity of MT

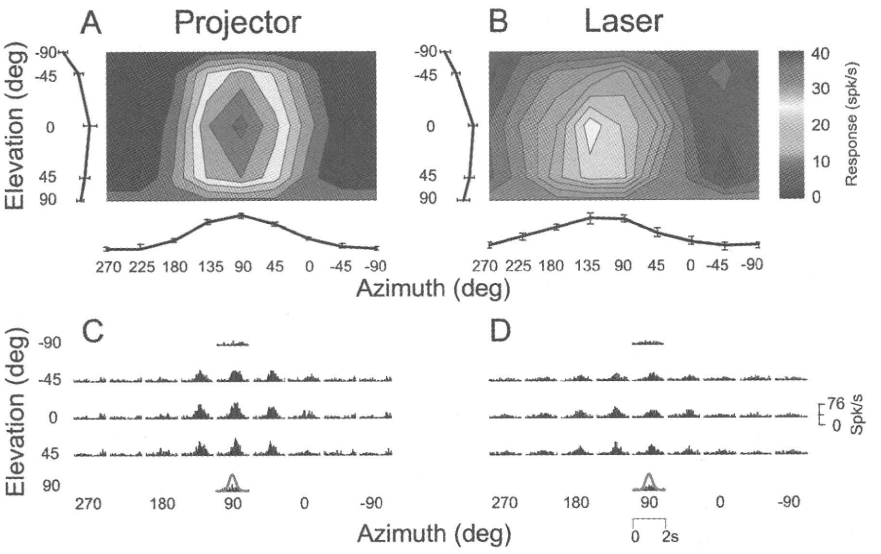

Figure 8. ***A–D***, 3D rotation tuning profiles for an MSTd neuron tested under the projector (***A***, ***C***) and laser (***B***, ***D***) conditions. Color contour maps in ***A*** and ***B*** show mean firing rate as a function of azimuth and elevation angles. PSTHs in ***C*** and ***D*** illustrate the corresponding temporal response profiles (format as in Fig. 1).

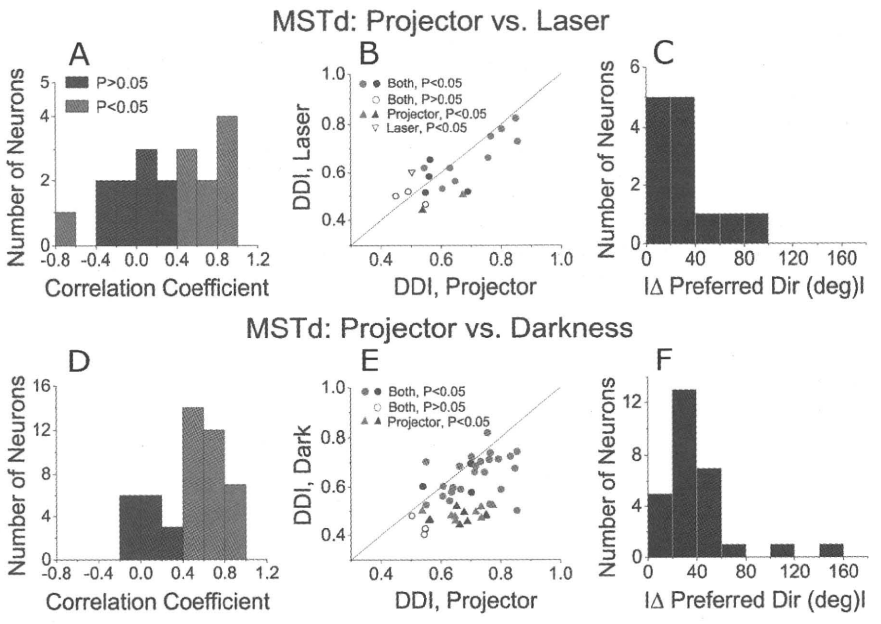

Figure 9. ***A–E***, Comparison of vestibular rotation tuning in area MSTd between the projector and laser (***A–C***) or darkness (***C–E***) conditions. ***A***, ***D***, Histograms show the distribution of correlation coefficients between 3D tuning profiles measured in the projector versus laser (***A***) or projector versus darkness (***D***) conditions. ***B***, ***E***, Scatter plots compare the DDI between stimulus conditions (format as in Figure 7). ***C***, ***F***, Histograms show the absolute difference in 3D direction preferences (|Δ preferred direction|) between projector and laser conditions (***C***) or projector and darkness conditions (***F***). Only cells with significant tuning in both conditions were included in these histograms.