

MSTd Neuronal Basis Functions for the Population Encoding of Heading Direction

S. Ben Hamed,¹ W. Page,^{1,2} C. Duffy,^{1,2} and A. Pouget¹

¹Department of Brain and Cognitive Science and the Center for Visual Science, University of Rochester, 14627; and ²Departments of Neurology, Neurobiology and Anatomy, and Ophthalmology, University of Rochester, School of Medicine and Dentistry, Rochester, New York 14642

Submitted 6 August 2002; accepted in final form 13 May 2003

Ben Hamed, S., W. Page, C. Duffy, and A. Pouget. MSTd neuronal basis functions for the population encoding of heading direction. *J Neurophysiol* 90: 549–558, 2003; 10.1152/jn.00639.2002. Basis functions have been extensively used in models of neural computation because they can be combined linearly to approximate any nonlinear functions of the encoded variables. We investigated whether dorsal medial superior temporal (MSTd) area neurons use basis functions to simultaneously encode heading direction, eye position, and the velocity of ocular pursuit. Using optimal linear estimators, we first show that the head-centered and eye-centered position of a focus of expansion (FOE) in optic flow, pursuit direction, and eye position can all be estimated from the single-trial responses of 144 MSTd neurons with an average accuracy of 2–3°, a value consistent with the discrimination thresholds measured in humans and monkeys. We then examined the format of the neural code for the head-centered position of the FOE, eye position, and pursuit direction. The basis function hypothesis predicts that a large majority of cells in MSTd should encode two or more signals simultaneously and combine these signals nonlinearly. Our analysis shows that 95% of the neurons encode two or more signals, whereas 76% code all three signals. Of the 95% of cells encoding two or more signals, 90% show nonlinear interactions between the encoded variables. These findings support the notion that MSTd may use basis functions to represent the FOE in optic flow, eye position, and pursuit.

INTRODUCTION

Our movements through the environment create the patterned visual motion of optic flow. During forward movement, this takes the form of an expanding flow field originating from a point in space known as the focus of expansion (FOE). The FOE in head-centered coordinates indicates heading direction and is an important cue for navigation (Gibson 1950). When the eyes are fixed and pointing straight ahead, the position of the FOE in head-centered coordinates coincides with its position in eye-centered coordinates. This is no longer the case when the eyes are moving, although the heading direction could still be recovered by combining the position of the FOE in eye-centered coordinates with signals such as eye position and velocity. This could account for the ability of humans and monkeys to discriminate heading direction with biases around 1–2° (Royden et al. 1992) and discrimination threshold within 1–3° (Britten and Van Wezel 2002; Warren and Hannon 1990).

Neurons in the medial superior temporal (MST) area of the macaque monkeys are believed to play a key role in the neural basis of these computations (Page and Duffy 1999). Page and Duffy (1999) showed that the position of the FOE can be decoded from the MST spike counts over a 1 s period averaged over 8 trials. These results showed that direction of the FOE, where direction refers to the angle of the FOE vector in polar coordinates, can be decoded with a bias of $\pm 10^\circ$ for FOE representation in head-centered coordinates during fixation and a bias of $\pm 16^\circ$ during pursuit (the bias is the difference between the mean estimate and the true value, also known as the average error). Moreover the population vector was found to increase with the eccentricity of the FOE. Bremmer and colleagues also reported that eye position can be read out from the averaged MST activity with a bias of 10^{-5} deg using an optimal linear estimator on averaged neuronal responses (Bremmer et al. 1998). These results demonstrate that the position of the FOE in head-centered coordinates and the position of the eyes are indeed encoded in area MST. However, they do not tell us how accurately this information is made available to the system on a real-time scale.

In this study, we first asked whether the position of the FOE can be decoded from the responses of MST neurons on *single trials*—as opposed to activity averaged over several trials. If so, how does the precision compare with behavioral performance (Britten and Van Wezel 2002; Royden et al. 1992; Warren and Hannon 1990)? Because perceptual decisions are necessarily based on single-trial activity, such a result would strengthen the notion that MST is indeed playing a central role in heading perception. To address this question, we split the data into two sets, one for training the decoder and one for testing it, a method known as cross validation. This method prevents overfitting and leads to accuracy measurements that generalize to new data sets (i.e., single-trial activity patterns from the same neuronal population) that were not used for either training or testing. We then used an optimal linear estimator of the Cartesian coordinates of the FOE, based on single-trial activity and using cross validation. We also trained similar optimal linear estimators for eye position and eye velocity, to determine whether these variables can also be read

Address for reprint requests: A. Pouget, Brain and Cognitive Science Department, Meliora Hall, University of Rochester, Rochester, NY 14627 (E-mail: alex@bcs.rochester.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked “advertisement” in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

out from single-trial activity and the accuracy to which this can be achieved.

To further characterize the neural code in MST, we also investigated its format because the format determines the computational properties of the code and how it might be used by downstream (i.e., higher) areas (Pouget and Snyder 2000). We considered the possibility that MSTd contains specialized neuronal populations for FOE, eye positions, and eye velocity and compared this model to one based on the nonlinear combination of two or more of these variables. The latter model would support the notion that MSTd uses a basis function representation (Poggio 1990; Pouget and Snyder 2000), a format that has several computational advantages. With such a representation, any function (i.e., any motor command), perceptual judgment, or behavior, depending on eye velocity, eye position, and the position of the FOE in eye-centered coordinates, could be computed by taking a simple linear combination of the responses of MST neurons (Poggio 1990; Pouget and Snyder 2000).

Moreover, basis function networks with lateral connections can implement a form of statistical inference known as maximum likelihood estimation, which allows these networks to perform optimal computations with population codes in the presence of uncertainty (Denève et al. 2001). Maximum likelihood estimation encompasses a large class of computations, including the estimation of heading direction independently of eye movements. It is therefore possible to relate the performance of human subjects in tasks involving judging heading direction, to the response of basis function neurons integrating the position of the FOE in eye-centered coordinates, as well as the position and velocity of the eyes. In summary, basis function representation offers several computational advantages and can be related to behavioral performance (Denève et al. 2001).

METHODS

Electrophysiological studies

The activity of 144 neurons from area MSTd, above the superior temporal sulcus (STS) (AP, -2 mm; ML, ± 15 mm), was recorded in five hemispheres of three macaque monkeys using standard unit neuronal recording methods, while the animals were performing a visual fixation task or a pursuit task (these data and the details of the experiment were previously presented in Page and Duffy 1999). Briefly, each monkey was seated in a primate chair in front of a $90^\circ \times 90^\circ$ tangent screen placed 48 cm away. During the fixation task, the monkeys had to maintain fixation at the center of the screen within a window of $\pm 3^\circ$ while a stimulus was presented for 1.5 to 2.5 s. On successful fixation, a sound reinforcement was given to the monkey as well as a liquid reward. During the pursuit task, the central fixation point was extinguished to reappear at 7.5° from the center in one of four possible directions (right, up, left, and down) for the trials carried out in the light and one of 8 possible directions (right, upper right, up, etc.) for the trials carried out in the dark. This galvanometer-animated pursuit target moved at $15^\circ/\text{s}$ across the center of the screen for 1 s and the monkey was required to follow it within a 3° range of accuracy. On successful pursuit, the animal received a liquid reward.

During both fixation and pursuit, the monkey could either be facing a blank screen except for the fixation (or pursuit) stimulus or a back-projected expanding dot pattern. This pattern could have 9 possible foci of expansion (FOE) presented in a pseudo-random order during a given block of fixation or pursuit (a central FOE and 8 FOEs distributed around the center at 45° intervals at an eccentricity of 30° ; an eccentricity of 15° was also used in the fixation task) and lasted for 1 s (see Page and Duffy 1999 for details).

Neuronal data were collected using standard electrophysiological single neuron recording techniques. For each task and each FOE condition, the neuronal response was either averaged over the entire 1 s period or binned in 100 ms bins. In the subsequent analyses on binned data, the first 100 ms data were not taken into account, given that this time range corresponds to the response latency of the neuron and to the establishment of its phasic response.

Decoding of MSTd activity

We used the activity of 144 MSTd neurons to estimate four variables: the position of the FOE in head- and eye-centered coordinates, the pursuit vector, and the position of the eyes. By activity we mean the spike counts over 1 s or 100 ms intervals depending on the signal being decoded. We used optimal linear estimators, that is, linear estimators whose weights are optimized to minimize the square distance between the estimate and true value. These estimators are formally equivalent to a fully connected two-layer network with the input layer containing 144 units, corresponding to the 144 MSTd neurons. The output layer contains 2 units, corresponding to the horizontal and vertical components of the estimated signals. These output units have linear activation functions; their activity is computed by taking the weighted sum of their inputs.

To optimize the weights, we trained the network using standard gradient descent techniques. In the case of the FOE, the data set consisted of 10,000 pairs of FOE positions and activity vectors used for training and 200 additional pairs used for testing, and generated as follows. For each neuron, the experimental data set contained 6–8 trials for each of the 9 FOE positions tested, corresponding to 8 positions in a circle of radius 30° plus the center. We used six of these trials for optimization and two for testing. For each FOE position, we generated about 1111 (10,000/9) activity patterns for training, by taking random combinations of the 6 trials set aside for training. This combination scheme allowed us to generate population patterns of activity for any stimulus configuration, even though the neurons were actually recorded one at a time, sometime on different days. Likewise, for each FOE, the 22 (200/9) activity patterns used for testing were generated by taking random combinations using the 2 trials reserved for testing. An additional 8 FOE positions in a circle of radius 15° had also been tested for 34 neurons and similar data sets were generated following the same procedure.

Once the data set was generated, we optimized the weights on the training data set and monitored performance on the testing set. The optimization was stopped when the performance on the testing started to decrease. This procedure prevents overfitting and provides an estimate of generalization performance (Morgan and Bourlard 1990). We then reiterated the whole procedure by splitting randomly the n available trials per neurons into $n - 2$ new trials for training and 2 for testing (where n was equal to 6 or 8). We performed five iterations altogether, for a total of $10,000 \times 5 = 50,000$ testing trials per position of the FOE. For each position of the FOE, we calculated the bias and SD of the estimate over the 50,000 trials (the bias was obtained for the x - and y -dimensions independently, by computing the absolute difference between the average estimate and the true position of the FOE). In the RESULTS section, we provide only the bias and SD averaged over all positions (9 positions in most cases). We averaged the biases and variances over all 9 positions to simplify the presentation of the results. There were no apparent systematic variations of those measures across space. The same exact procedure was followed for estimating pursuit, eye position, and the position of the FOE in eye-centered coordinates.

For pursuit, the experimental data corresponded to a 1 s period of activity for 5 different values of pursuit ($[0^\circ, 0^\circ]$, $[15^\circ, 0^\circ]$, $[0^\circ, 15^\circ]$, $[-15^\circ, 0^\circ]$, $[0^\circ, -15^\circ]$): for eye position and the position of the FOE in eye-centered coordinates, the experimental data had to be preprocessed because neither of these variables was maintained constant over a 1 s duration on trials in which the eyes were pursuing.

Specifically, we divided each trial into ten 100 ms time bins and considered eye position and the position of the FOE in eye-centered coordinates as constant over those time bins.

For the sake of comparison with previous studies, we also trained one network on decoding the FOE in head-centered coordinates based on the activity of MSTd neurons, averaged over all trials (8 per condition). In this case, the weights were optimized over 100,000 presentations of the average activity vector.

Invariance and nonlinearity

Effects of FOE and pursuit direction on the neurons' response were assessed by an ANOVA using the Matlab statistical toolbox. A 2-way, between groups design was used to measure FOE and pursuit main effects as well as FOE \times pursuit interaction effects from data averaged on 1 s. The same analysis was also used to measure FOE and eye position main effects as well as FOE \times eye position interaction effects from data averaged on 100 ms. All results were tested for a statistical significance at the $P = 0.01$ level of confidence. Neuronal activities were log-transformed to decrease differences in the response SD between each condition.

A cell was deemed to be invariant to a particular variable if the ANOVA showed a significant modulation to this variable but to no other. Thus cells invariant to pursuit (to FOE) were defined as those cells being significantly modulated by FOE (with respect to pursuit) but not by pursuit (with respect to FOE) or eye position. For the cells combining 2 or more signals, we also tested whether these signals were combined nonlinearly by checking whether the ANOVA also revealed a significant interaction between these two parameters.

RESULTS

We analyzed the responses of 144 neurons that were recorded in dorsal MST (MSTd) in three behaving rhesus monkeys. These data constituted the basis of a previous report of neurophysiological findings (Page and Duffy 1999). We used optimal linear estimators to estimate the position of the FOE in head- and eye-centered coordinates, as well as the velocity and position of the eyes.

Deriving of FOE and pursuit direction from averaged MST neuronal responses

We first performed an analysis based on the *averaged* activity of MSTd neurons to compare our results to previous studies. Neuronal activity was averaged over all trials with identical FOEs, both while the monkeys were fixating a central point and while they pursued a target moving in one of four directions. This activity was used to decode the position of the FOE in head-centered coordinates. The distance of the network response from the true FOE was on average: horizontal = $0.003 \pm 0.002^\circ$, vertical = $0.0031 \pm 0.0022^\circ$ (the means and SDs are computed over the 9 positions of the FOE tested). In polar coordinates (i.e., radius and angle as opposed to vertical and horizontal positions), the FOE position can be estimated with an error of $0.010 \pm 0.010^\circ$. This performance measure was determined from averaged neuronal responses irrespective of eye velocity, including null eye velocity (i.e., fixation). We performed this transformation because a previous attempt to decode the direction of the FOE from the average response of MSTd neurons using a population vector estimator had reported an error of $9.6 \pm 6.7^\circ$ on data collected during fixation and $16.0 \pm 10.6^\circ$ on data collected during pursuit. Therefore the optimal linear estimator is more precise than the population

vector approach by a factor of 10^3 . Another benefit is that the optimal linear estimator can estimate an FOE at the center of the screen.

Deriving FOE and pursuit direction from single-trial neuronal responses

The result that the position of the FOE in head-centered coordinates can be recovered with a bias of 0.01° is a strong indication that MSTd encodes this variable. However, this value was obtained on averaged responses, using identical data sets for training and testing. To compare the accuracy of the neural code to behavioral discrimination thresholds, we need to obtain a generalization performance. This requires estimating the position of the FOE from single-trial activity patterns that were not used for training the estimator. The use of single-trial data are key here: a subject involved in a psychophysical task does not have the option of averaging the data over many trials as we did in the previous section. Accordingly, we generated single-trial data as described in the METHODS section and we divided the data set into training and testing trials. The estimator was optimized on the training set and performance was assessed on single trials from the testing set (see METHODS). This is the procedure we follow in the remainder of the study.

FOE IN HEAD-CENTERED COORDINATES AND PURSUIT IN THE LIGHT. Spike counts from single trial neuronal responses were cumulated over the 1 s duration of the visual stimuli and pursuit directions. Nine FOEs were presented, the central FOE and 8 others at 30° of eccentricity distributed around the center at 45° intervals. Five pursuit conditions were used, fixation and the four cardinal directions on the horizontal and vertical axes. All results from now on will be presented in the form of $A^\circ \pm B^\circ$, where A° corresponds to the bias (the difference between the mean of the estimate and the true value), whereas B° is the SD of the estimate, both computed over all testing trials, and then averaged across all positions (see METHODS for details).

The single-trial estimates of FOE and pursuit direction in the light are more accurate than those obtained by the population vector approach from the averaged activity: the position of the FOE in head-centered coordinates is recovered with an error of ($1.57 \pm 3.04^\circ$, $0.63 \pm 2.39^\circ$) from its actual location and the pursuit vector is recovered with similar error ($1.54 \pm 3.94^\circ$, $1.2 \pm 3.19^\circ$) (Fig. 1).

PURSUIT DIRECTION IN THE DARK. Single trial neuronal responses were also obtained during pursuit in the dark (fixation and 8 pursuit directions). A linear estimator was specifically optimized to process responses from this experimental condition. We found both lower biases and smaller SDs in the dark than in the light with an error for estimating the pursuit direction in the dark of ($1.26 \pm 1.52^\circ$, $1 \pm 1.39^\circ$). This suggests that the presence of optic flow introduces additional variability in the pursuit responses of MSTd neurons.

FOE IN HEAD-CENTERED COORDINATES WITH 15° AND 30° ECCENTRICITIES. Thirty-four neurons were studied with 17 FOE stimuli, one in the center, 8 with an eccentricity of 15° , and 8 with an eccentricity of 30° . We trained an optimal linear estimator to decode the activity of these 34 neurons from single trial neuronal responses averaged over the 1 s duration of these visual stimuli presented during fixation. The average error for the 8 FOEs at 30° was ($7.79 \pm 5.98^\circ$, $7.97 \pm 5.12^\circ$), and for

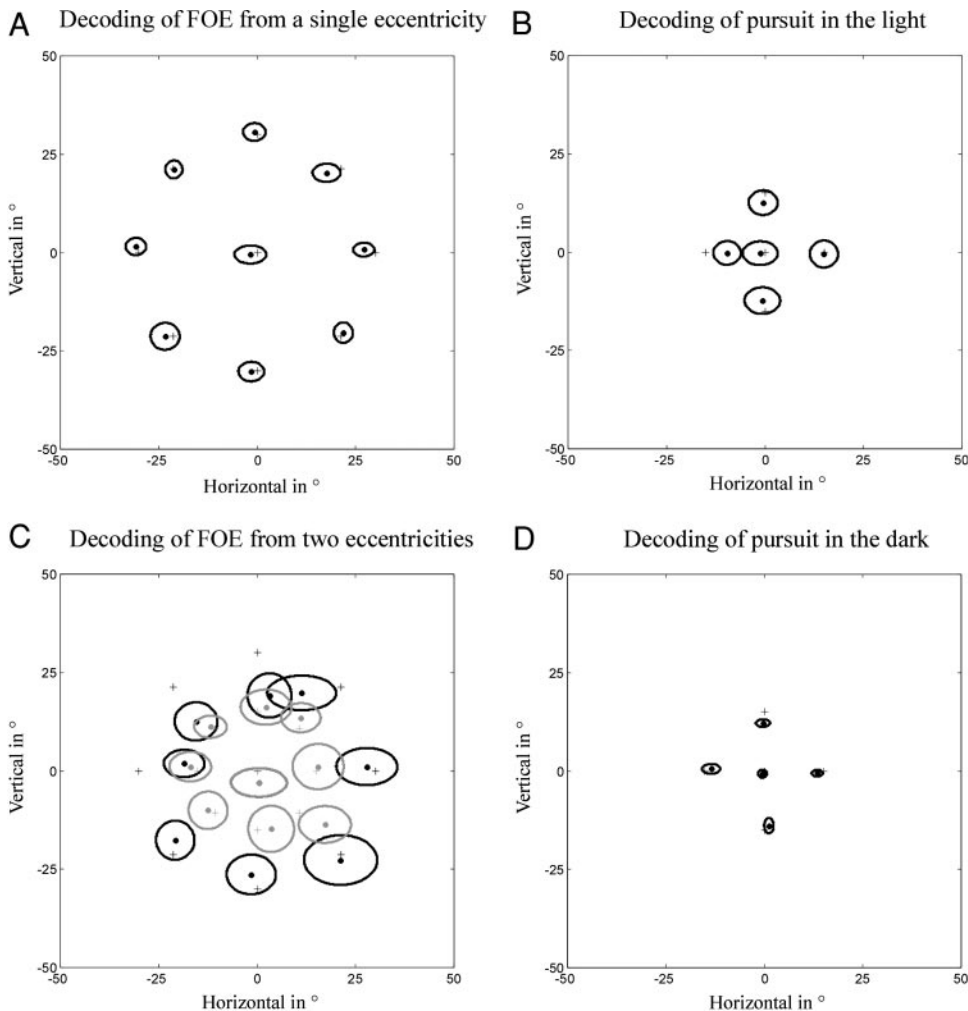


FIG. 1. Decoding of focus of expansion (FOE) in head-centered coordinates and of pursuit direction. The network is trained on successively and randomly selected neuronal responses to FOE presentation or pursuit direction during individual trials and is tested on an independent testing data set (see METHODS section). All analyses were carried out on a set of 144 cells except for C, which was carried out on a subset of 34 cells. Actual location of FOEs and pursuit directions are indicated by a + sign. For each analysis are represented mean estimate (dot) and SD around that mean (ellipse). Standard errors do not appear on graphs because of their very small values. *A*: decoding FOE after training on trials presented at a single eccentricity of 30° from the center of the visual field. Mean error in the estimates was $(1.575 \pm 3.04^\circ, 0.627 \pm 2.39^\circ)$. *B*: decoding of pursuit direction in the light. Mean error in the estimates was $(1.392 \pm 3.69^\circ, 1.378 \pm 3.02^\circ)$. *C*: decoding of FOE after training on trials presented at 30° (in black) and 15° (in gray) eccentricities. Mean error in the estimates was $(7.79 \pm 5.98^\circ, 7.97 \pm 5.12^\circ)$ when decoding trials tested for FOE at 30° of eccentricity and $(3.62 \pm 6.78^\circ, 3.87 \pm 4.96^\circ)$ when decoding trials tested for an FOE at 15° of eccentricity. *D*: decoding of pursuit direction in the dark. Mean error in estimates was $(1.129 \pm 1.31^\circ, 0.979 \pm 1.08^\circ)$.

the 8 FOEs at 15° , $(3.62 \pm 6.78^\circ, 3.87 \pm 4.96^\circ)$ (Fig. 1C). The error at 30° is greater than that found previously because we used 34 instead of 144 neurons. Note also that the bias at 15° is half that obtained at 30° . This does not mean that MSTd neurons encode the position of the FOE for small eccentricities better than for large ones because the performance of a linear estimator is not always directly proportional to the information content. Consider for instance a population code in which neurons have equally spaced bell-shaped tuning curves to a variable, θ . Ignoring edge effects, the discrimination threshold of an ideal observer in this case is the same for all θ , from which it follows that Fisher information is constant [because Fisher information is inversely proportional to the discrimination threshold of an ideal observer (Seung 1993)]. This is not what would be obtained with a linear estimator. An observer performing discrimination with a linear estimator would end up with a performance proportional to θ (Baldi and Heiligenberg 1988; Pouget and Sejnowski 1994).

In other words, one would have to use Fisher or Shannon information to assess more precisely whether the code in MSTd is indeed less precise in the periphery. Unfortunately, either of these methods requires a precise assessment of the joint probability distribution over the activity of all neurons for all stimuli. Even if we assume that the neuronal responses are independent and follow Gaussian distributions, we would still need to estimate the mean and variance of the firing rate of

each neuron for each experimental condition. Unfortunately, with a maximum of 8 trials per condition, we could not obtain reliable estimates of these parameters. Nevertheless, our results using optimal linear estimators suggest that 34 neurons are sufficient to estimate the position of the FOE with a bias of at most 8° and SD of at most 6° within a 30° circle.

EYE POSITION AND EYE-CENTERED FOE. These two signals were decoded from single trial neuronal responses averaged over ten 100-ms bins encompassing the total 1 s duration of the visual stimuli and pursuit eye movements.

Eye position was recovered with an error of $(1.51 \pm 1.80^\circ, 0.71 \pm 3.29^\circ)$ (Fig. 2A). This cannot be directly compared with FOE and pursuit estimators based on 1 s samples of activity. However, assuming temporal invariance of response statistics, we would expect the estimation of eye position over a 1 s interval to improve by a factor of the square root of 10, 10 being the number of 100 ms bins in a 1 s interval. This leads to a predicted resolution of $(0.48 \pm 0.57^\circ, 0.23 \pm 1.04^\circ)$, 3 to 6 times smaller than those obtained for FOE and pursuit. There are two possible explanations for this observation: 1) the mutual information between neural activity in MSTd and eye position is higher than that for any other variables, and 2) the format of the code is better adapted to a linear readout. Both of these explanations might apply but we cannot determine their respective contributions without a direct measure of mutual information.

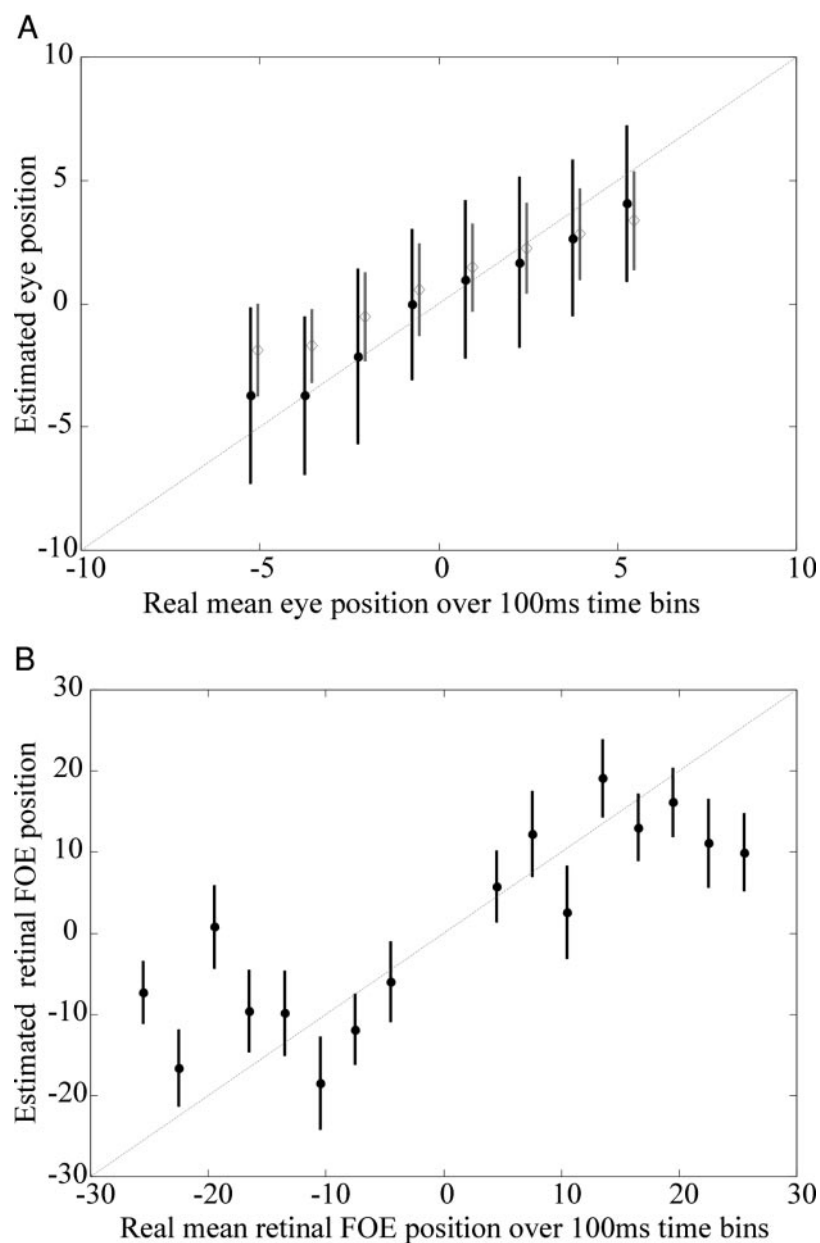


FIG. 2. Decoding of eye position and of FOE in eye-centered coordinates over 100 ms time bins (see METHODS section). For each analysis are represented mean estimate (dot) and SD around that mean (vertical lines). Standard errors do not appear on graphs because their values do not exceed diameter of the dots. *A*: decoding of horizontal (in gray) and vertical (in black) eye position. Mean error in the estimates was $(1.515 \pm 1.80^\circ, 0.714 \pm 3.29^\circ)$. *B*: decoding of FOE in eye-centered coordinates, for trials recorded during upward or downward pursuit, whereas FOE was present in center of screen. Mean error in the estimates calculated for FOEs located in center of screen, up or down, was $7.308 \pm 4.89^\circ$.

In the case of eye-centered FOE coordinates, we found an error of $(7.31 \pm 4.89^\circ, 7.16 \pm 5.2^\circ, \text{Fig. } 2B)$. Applying the same correction used for eye position, we find that the eye-centered FOE can be estimated within $(2.3 \pm 1.55^\circ, 2.26 \pm 1.64^\circ)$. These numbers are in the same range as those obtained for FOE in head-centered coordinates and pursuit.

Basis functions for heading direction

In the previous section, we showed that several signals are present in MSTd and can be decoded to a very high degree of precision from the response of the population of neurons. In this section, we investigate whether MSTd neuronal responses can serve as basis functions for the processing of heading direction. Two necessary conditions must be met (Pouget and Sejnowski 1997): 1) a large proportion of MSTd neurons should be selective to two or more signals; that is, it should not be the case that most neurons in MST are specialized for one

signal only; and 2) neurons selective to two or more variables should show a nonlinear tuning to those variables.

ONE POPULATION OR SEVERAL SUBPOPULATIONS? The results presented above show that several signals are present in MSTd area at the level of the population and that these signals can be recovered at a very high degree of precision. However, we do not know whether these signals are encoded by a single population or by several subpopulations.

A 2-way ANOVA of single-trial firing rates with main factors of FOE and pursuit show that the vast majority of cells are tuned for both FOE stimuli and pursuit direction (126/144, 87.5%; $P < 0.01$) (Fig. 3A). An additional 3.5% (5/144, $P < 0.01$) show FOE selectivity alone and 7.6% (11/144, $P < .01$) show pursuit selectivity alone. A similar analysis on cell activity averaged over 100 ms bins indicates that 16/144 (11.1%, $P < 0.01$) of the cells are responsive to FOE but not to eye position, and 28/144 (19.5%, $P < 0.01$) have the opposite

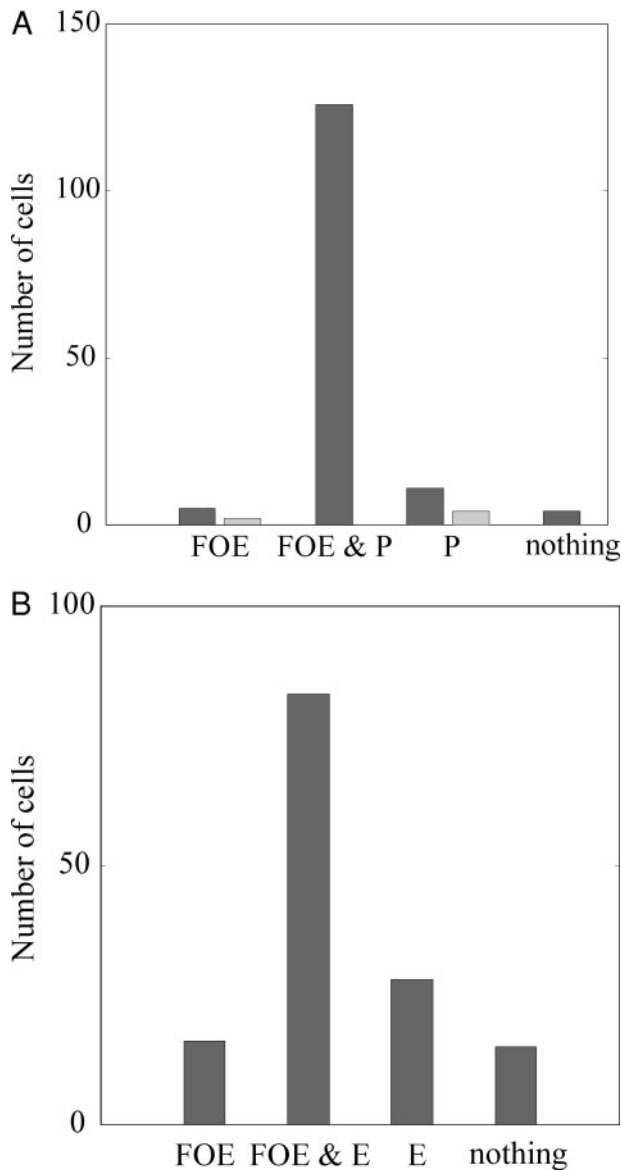


FIG. 3. Distribution of FOE, pursuit, and eye position selectivities. *A*: in black, number of cells selectively responsive to FOE (5/144, 3.5%), to pursuit (11/144, 7.6%), or to both factors (126/144, 87.5%). In white, number of cells that respond to FOE (resp. pursuit) but whose response is not significantly affected by eye position (2/144, 1.4%, resp. 4/144, 2.8%). *B*: number of cells selectively responsive to eye position (28, 19.5%), to FOE (16, 11.1%), or to both (83, 57.6%).

trend, whereas again, more than half of the cells are tuned to both factors (83/144, 57.6%; $P < 0.01$) (Fig. 3*B*). Interestingly, when all three factors are taken together, only 6 cells are found to be selective to one variable only (Fig. 3*A*), whereas 104 (72.2%) cells are determined to be selective to all of FOE, pursuit, and eye position. Some 95.8% (138/144) of MSTd cells thus encode 2 or more variables. This is likely to be a lower bound for the percentage of MSTd cells encoding 2 or more signals, given that only a limited number of FOEs, pursuit vectors, and eye positions were tested (9 FOEs, 4 pursuit vectors, 40 eye positions) with only a limited number of trials per condition (8 trials).

We also tested MSTd neuronal encoding of two or more variables using our optimal linear estimator of FOE position in head-centered coordinates and pursuit direction. An optimal linear

estimator is equivalent to a two-layer network. Each input unit, corresponding to one MSTd neuron, sends 4 weights to the 4 output units, the horizontal and vertical positions of the FOE, and horizontal and vertical components of the pursuit vector. Figure 4*a* shows the distribution of the weights for 2 of those 4 components, that is, the horizontal FOE and pursuit. The weights are not clustered around the vertical and horizontal axes, as would be expected if most neurons were specialized for only one variable. Instead, they are homogeneously scattered in the two-dimensional (2D) plane, demonstrating that most neurons contribute to the estimates of both variables. A similar pattern is obtained when plotting the weights for the horizontal FOE as a function of the weights for horizontal eye position (Fig. 4*B*). Here again, the weights of these cells are uniformly scattered in the 2D plane with no clustering along the FOE and eye position axes.

LINEAR OR NONLINEAR INPUT INTERACTIONS? Cortical neurons are inherently nonlinear because their responses saturate at a minimum rate of 0 Hz and they rarely fire beyond 200 Hz (at least for excitatory neurons). Therefore the important question is not so much whether neurons are nonlinear but whether they show these nonlinearities within the range of inputs typically encountered. For instance, it has been claimed that neurons in the primary motor cortex have a cosine tuning to the direction of hand movements (Georgopoulos et al. 1986). This would suggest that these neurons compute a dot product between their preferred direction and the direction of the hand movement, in which case their responses are linear over the entire range of movement directions.

To investigate this issue, we first performed linear regressions over the activity of each neuron as a function of the position of the FOE and the pursuit vector. We found that the average $R^2 = 0.30 \pm 0.20$, meaning that the linear hypothesis accounts for only 30% of the variance of the cells. These numbers are particularly low given that only 9 positions of the FOE in the 2D plane and 5 pursuit vectors were tested in the experiment. With such a sparse sampling, it is typically difficult to detect nonlinear trends and, yet, they appear to be present within the range of input tested. However, low R^2 might be attributable to other factors beside nonlinear interactions, such as high noise level, and hence the need for other tests.

Second, we used a 2-way ANOVA to analyze interaction effects between FOE and pursuit on the response of neurons significantly modulated by both factors. We ran this analysis because the interaction term in an ANOVA is a test of nonlinearity. For instance, if a neuron receives inputs X and Y , and computes the product XY (which is a nonlinear combination of X and Y), the interaction should be significant, assuming the noise is low enough. Our results showed that 98/126 (77.8%) of the cells show a nonlinear interaction between FOE and pursuit (Fig. 5). A similar approach on FOE and eye position modulation reveals that a similar percentage of cells (66/83, 79.5%) shows a nonlinear interaction between FOE and eye position. By combining both analyses, we found that an additional 14 cells (11.1%) of the 28 (126 - 98) cells that did not show a nonlinear interaction between FOE and pursuit show a nonlinear interaction between FOE and eye position (Fig. 5). Thus FOE and pursuit or eye position interact nonlinearly in $\geq 112/126$ (88.9%) cells responsive to both FOE and pursuit.

Third, we examined nonlinear interactions by testing the ability of the optimal linear estimator to generalize across data

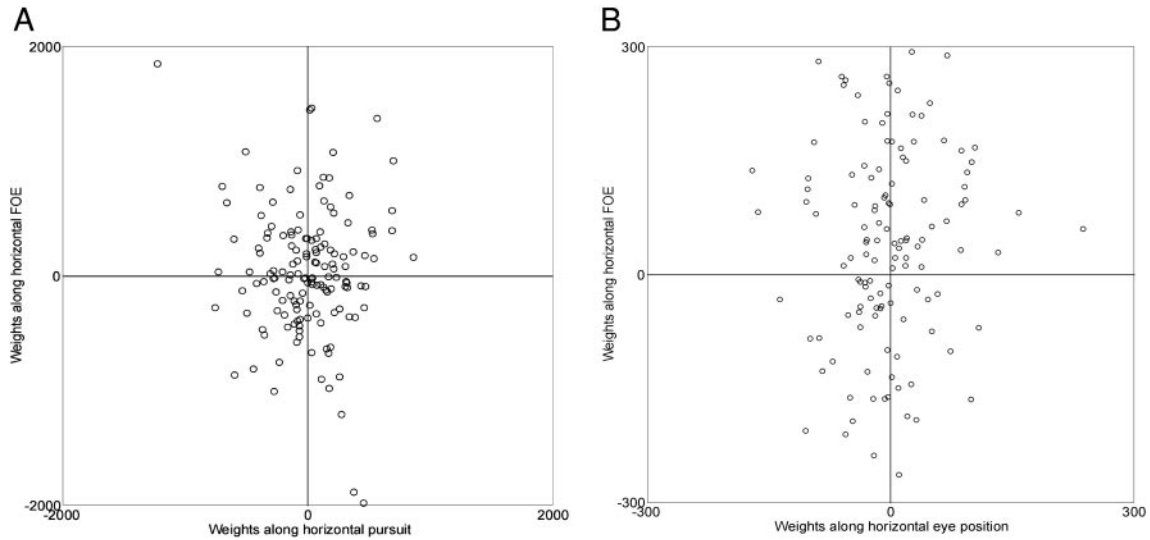


FIG. 4. Analysis of weights developed by networks when decoding FOE, pursuit, and eye position. *A*: weights developed when decoding horizontal FOE as function of weights developed while decoding the horizontal dimension of pursuit. *B*: weights developed when decoding horizontal FOE as function of weights developed while decoding the horizontal dimension of eye position.

sets. Let us call $\mathbf{W}_{p,light}$ the weights of the optimal linear estimator of pursuit, $\hat{\mathbf{P}}_{light}$, obtained from the response of the neurons in conditions involving a combination of FOE and pursuit, and $\hat{\mathbf{P}}_{dark}$ the estimator of pursuit obtained by applying $\mathbf{W}_{p,light}$ to the neural activity recorded in the dark. As shown in the APPENDIX, under the assumption of linear interactions, the difference between the conditional mean of $\hat{\mathbf{P}}_{light}$ and the conditional mean of $\hat{\mathbf{P}}_{dark}$ should be a constant for all pursuit directions.

Thus we first obtained $\mathbf{W}_{p,light}$ by training a linear estimator

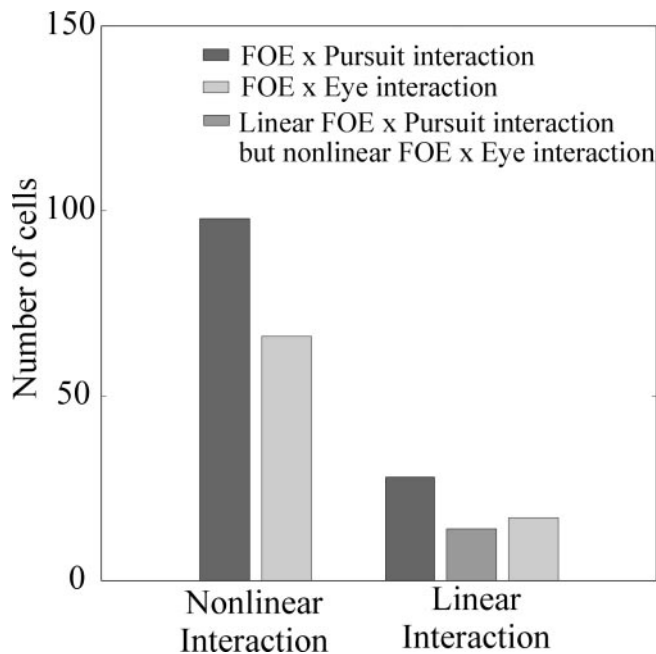


FIG. 5. Interaction of FOE with pursuit signals and FOE with eye position signals at single cell level. In dark gray, number of cells whose responses have linear interaction (resp. in light gray, nonlinear interactions) between FOE and pursuit signals (28; resp. 98) and FOE and eye position signals (17; resp. 66) at signal cell level. In intermediate gray is plotted number of cells showing linear interaction between FOE and pursuit and nonlinear interaction between FOE and eye position (14).

to decode pursuit from the neuronal activity of MSTd recorded during pursuit eye movements in the presence of expansion patterns on the screen. Black labels in Fig. 6 show the conditional means ($\mathbf{E}[\hat{\mathbf{P}}_{light}|\mathbf{P}]$) and SD of the estimates of pursuit after optimal training. We then computed $\hat{\mathbf{P}}_{dark}$, that is, the estimate of pursuit obtained by applying $\mathbf{W}_{p,light}$ to the activity of the neurons recorded during pursuit movement in the ab-

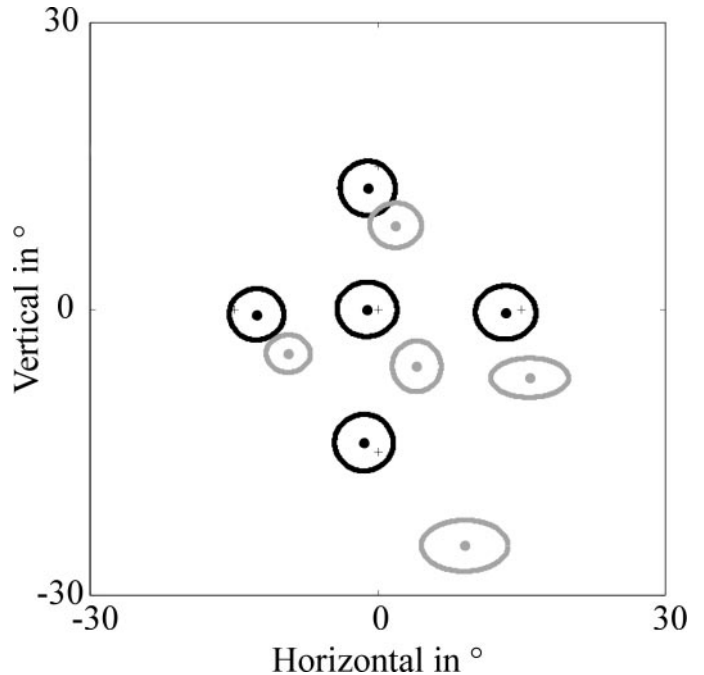


FIG. 6. Pursuit in the dark and pursuit in presence of expansion patterns cannot be decoded by the same set of weights. Crosses indicate end point of actual pursuit vector the monkey was asked to perform. In black is result of decoding pursuit in presence of different expansion patterns. Central black dots represent mean end point of network on testing iterations. Black circles represent scatter (taken as SD) of estimates around mean value. In gray is result of decoding pursuit in the dark using the weights obtained when training network to decode pursuit in the light. Dots and circles same as previously. Standard errors do not appear because of their very small values.

sence of expanding flow fields. Gray labels in Fig. 6 show the conditional means ($\mathbf{E}[\hat{\mathbf{P}}_{\text{dark}}|\mathbf{P}]$) and SDs of the estimates of pursuit in this condition.

A minimum of 2215 estimates is available per pursuit direction and per luminosity condition (dark versus light). Two-way ANOVAs with the x- and the y-coordinates of the estimates as dependent factors, and pursuit condition and luminosity condition as independent factors reveal that, in both cases, the estimates in the dark are significantly different from those in the light ($P < 0.001$) and this difference is not systematic across all pursuit conditions (a significant interaction effect is noted, $P < 0.001$). Thus the difference between the two estimates $\hat{\mathbf{P}}_{\text{light}}$ and $\hat{\mathbf{P}}_{\text{dark}}$ depends on pursuit direction and is not constant. This strongly suggests that the neurons combine their selectivities to **FOE** and **P** in a nonlinear way.

DISCUSSION

Coding accuracy in MST

We have demonstrated that MSTd encodes the position of the FOE in head- and eye-centered coordinates, as well as the position and velocity of the eyes, with a high degree of accuracy on single trials. In the case of the FOE, its position in head-centered coordinates could be recovered with a bias of 0.5–1.5° and SD of 2.4–3°. The value of the bias is similar to the biases of 1–2° found in humans by Royden et al. (1992) during pursuit eye movements.

To relate these results to behavioral discrimination thresholds, we need to assess the discrimination thresholds of our optimal linear estimator. In the experiments of Warren and Hannon (1990) in humans, and those of Britten and Van Wezel (2002) in monkeys, the thresholds were obtained by asking subjects to determine whether the FOE was left or right of a fixed reference target. Thresholds were defined as the difference between the positions of the FOE and the reference corresponding to a performance of 84% correct responses. For such experiments, the discrimination thresholds ($\delta\theta$) of our optimal linear estimator would be given by $\delta\theta = \sigma/b'$, where σ is the SD of the estimate and b' is the derivative of the bias. Note that the bias by itself is irrelevant for discrimination thresholds; only the slope (or derivative) matters. We have found that the derivative of the bias is 0.98 on average for our optimal linear estimator of the position of the FOE in head-centered coordinates during pursuit. Therefore the discrimination thresholds are within the range of 2.45–3.06°, which is consistent with the range reported in humans, 1–3° (Warren and Hannon 1990), and monkeys, 1.5–3° (Britten and Van Wezel 2002). These results provide further evidence that MST may be involved in heading perception (Duffy 1998; Duffy and Wurtz 1995), although they do not establish a causal relationship between the perception of heading direction and the neural activity in MSTd.

Nevertheless, there are several limitations to keep in mind. First, the optical flow fields used in the behavioral and neurophysiological experiments are not identical, even though they both contained foci of expansion. Second, the neuronal responses we used were collected with single electrodes such that correlations among MSTd neurons were not present in our data. We are referring here to correlations between pairs of neurons over several trials for the same stimulus (known as

noise correlations). It is possible that the inclusion of these correlations would significantly change the accuracy we observed. For instance, Zohary and Newsome (1994) showed that the correlations measured in area MT tend to increase the discrimination threshold of an ideal observer compared with a situation in which the neurons are independent. This effect, however, is small for samples on the order of 100 neurons, as in our sample of 146 neurons.

Third, it is important to remember that all our results are based on optimal linear estimators. This approach was chosen because of the small number of trials available per conditions in our data set. However, although an optimal linear decoder is optimal among all linear decoders, it is not *the* optimal decoder. As a result, it provides only a lower bound on the information content of a particular code. The actual information content could be much larger and this is particularly true if MSTd uses basis functions in the form of nonlinear combinations of Gaussian tuning curves. In this case, the optimal decoder is the maximum likelihood estimator; and its performance directly reflects the amount of Fisher information available in the neural activities (Pouget et al. 1998; Seung and Sompolinsky 1993). Therefore it is possible that the position of the FOE, pursuit vector, and eye position can be estimated with higher accuracies than found behaviorally. This difference might reflect noise entering the system between MSTd and other cortical areas involved in perceptual decisions (Shadlen et al. 1996). It could also be attributed to the fact that the nervous system uses a suboptimal readout algorithm.

Finally, it is important to keep in mind that the data we used were collected in response to motion flow fields with radial symmetry and one depth plane. Motion flow fields are more complex for curved self-movements and multiple depth planes. These stimulus parameters are known to influence MSTd neuronal responses (Duffy and Wurtz 1997; Upadhyay et al. 2000). Future experiments may determine whether our linear decoding scheme also works for these more complex flow fields.

Basis functions in MSTd

We found no evidence that the encoding of FOE, eye position, and eye velocity involves distinct subpopulations of MSTd neurons. Instead, the majority of the neurons encode nonlinear combinations of two or more of these signals. These findings suggest that MST neurons compute basis functions of their inputs, a scheme known to be computationally efficient and robust to noise (Pouget and Snyder 2000).

Our results also show that approximately 90% of MSTd neurons mix nonlinearly two or more signals, whereas there is little evidence for specialized populations in MSTd (specialized in the sense of encoding a single signal). This does not support the claim that $\leq 26\%$ of the neurons in MSTd are invariant to the position of the FOE in head-centered coordinates (Bradley et al. 1996). As suggested previously (Page and Duffy 1999), this result reflects the fact that in this earlier study, FOE selectivity was tested only along each neuron's preferred pursuit axis. When the test is extended to two or more pursuit axes, few if any neurons are invariant to the position of the FOE in head-centered coordinates.

The nonlinear mixing of two or more signals by 90% of MSTd neurons suggests that MSTd could use basis functions to

co-represent various types of information. This is a computationally robust strategy (Pouget and Snyder 2000) that would allow downstream neurons that receive input from MSTd, to compute a variety of nonlinear functions of FOE position, eye position, and eye velocity by linearly combining the activities of MSTd neurons. MSTd's projections to other areas might then control distinct behaviors, such as navigation or eye movements. This versatility would be lost if the neurons projecting out of MSTd were exclusively encoding one or the other type of information.

Consistent with basis function encoding of FOE, pursuit (**P**), and eye position (**E**) in MSTd is the finding that all three variables were accurately decoded by the optimal linear estimator. Indeed, estimating these three input variables to MSTd is equivalent to analyzing them without further transformation [i.e., to computing a particular function of FOE, P, and E, i.e., the identity function, $f(\text{FOE}, \mathbf{P}, \mathbf{E}) = (\text{FOE}, \mathbf{P}, \mathbf{E})$]. This is a particularly simple function, but a function nonetheless, and as such, it can be obtained by taking a linear combination of basis functions of **FOE**, **P**, and **E**. This is precisely what is done with a linear estimator.

However, we do not claim that the nervous system uses optimal linear estimators to decode MSTd activity. Our claim concerns the nature of the neural representation in MSTd, not how it is read out. We must await additional data to determine how MSTd activity is read out by downstream neurons. Moreover, linear decoders are not particularly good choices from a statistical point of view, because they tend to be suboptimal. A better option is to use a maximum likelihood estimator, which typically outperforms linear estimators (Seung and Sompolinsky 1993). Interestingly, Denève et al. (1999) showed that it is possible to turn a basis function network into a maximum likelihood estimator by adding recurrent connections and nonlinear activation functions to all neurons. It is too early to tell whether the cortex, and in particular MSTd, contains such a global architecture.

We are not the first investigators to use the basis function framework to characterize a particular cortical representation; in fact, this representational scheme has been considered in many other areas along the ventral or dorsal pathways of the visual system (Logothetis and Pauls 1995; Poggio 1990; Poggio and Girosi 1990; Pouget and Sejnowski 1994, 1997). To our knowledge, however, this is the first demonstration based on single trial responses that cortical neurons convey nonlinear mixtures of multiple signals in a way consistent with basis functions.

Relation to previous models

There are strong similarities between the basis function representational scheme we propose and previous models of heading direction. Beintema and van den Berg (1998) proposed that the computation of heading direction involves neurons whose receptive fields act as motion templates multiplied by eye velocity. As shown by Pouget and Sejnowski (1997) in the context of coordinate transforms in the parietal lobe, such template units are equivalent to basis functions. The advantage of using the term "basis function" is that it emphasizes the generality of the approach. Basis functions provide a robust solution for approximating nonlinear functions (Poggio 1990), and, as such, can be applied to a large variety of problems, including, but not limited to, the computation of heading direction.

Moreover, thinking in terms of basis functions makes it easier to identify critical aspects of MSTd neuronal responses. For instance, one may ask whether a precise multiplication is required for the model of Beintema and van den Berg to work. If the answer were to be yes, this would be problematic because the data of Page and Duffy (2003) show that MSTd neurons are not perfect multipliers. Fortunately, precise multiplication is not required to preserve the computational properties of basis functions. All that is needed is the capacity to combine several signals nonlinearly, as described in this report.

Therefore the picture that emerges from our work is a variation of what Beintema and van der Berg propose, in which the basis functions do not simply multiply the motion evoked response with the eye velocity signals, but perform more complicated nonlinear combinations, leading, in some cases, to shifts in preferred FOEs. The recent model of Denève et al. (2001) shows how such basis functions can emerge in a recurrent network model that behaves like a Bayesian estimator.

Interestingly, the model of Perrone and Stone (1994) can also be related to the basis function framework. In their case, they investigated neural solutions to the computation of heading independently of eye movement on the basis of retinal cues only. This problem boils down to a nonlinear function approximation, in which the input is the motion flow field and the output is heading direction. The intermediate stage suggested by Perrone and Stone involves template units optimally tuned to one particular motion flow field. These template units are very similar to the view based units used by Edelman and Poggio (1990, 1991) in their radial basis function model of object recognition. Therefore the models of Perrone and Stone (1994) and of Beintema and van den Berg (1998) rely on similar computational strategies. The difference lies in the cues they are considering: Perrone and Stone rely on motion flow fields containing multiple depth planes, whereas Beintema and van den Berg use an extraretinal signal encoding eye velocity. Whether Perrone and Stone's basis functions are present in MSTd cannot be assessed with our data set because only one depth plane was used. Likewise, we cannot test the validity of "decomposition models" (Heeger and Jepson 1992; Hildreth 1992; Rieger and Lawton 1985; Royden 1997) with our analysis because, like the model of Perrone and Stone, decomposition models can be tested only with motion flow fields involving multiple depth planes.

Basis functions and perceptual invariance

The use of basis functions has implications for our understanding of the neural basis of perceptual invariances. It is often assumed that the neural correlate of perceptual invariances involves neurons whose tuning properties show the same invariance. Thus it has been proposed that our ability to judge heading direction accurately in the presence of pursuit eye movements is related to the existence of neurons whose tuning to the position of the FOE in head-centered coordinates is invariant across different eye positions and velocities (Bradley et al. 1996). Page and Duffy's (1999) results, and the analysis presented here, do not support Bradley and Andersen's assertion that there are head-centered neurons in MST, although such neurons might exist elsewhere. Because the heading direction in head-centered coordinates can be estimated from the response of MSTd neurons with a simple linear estimator with accuracy similar to the one measured behaviorally, any area

involved in perceptual judgment and receiving the activity of MSTd neurons can readily recover this information (see also Siegel 1998). Given that evidence for basis functions have been found in a wide variety of cortical areas, it is likely that this conclusion could extend to other functional domains.

APPENDIX

The aim of this analysis is to test for nonlinear interactions using optimal linear estimators. Let us consider the linear assumption and let $g_i(\text{FOE})$ and $h_i(\mathbf{P})$ be the response function of a neuron indexed by the letter i , to **FOE** and **P** (where g_i and h_i can be linear or nonlinear functions). The overall response function of neuron i to **FOE** and **P**, denoted r_i , is the sum of its selectivity to both variables, plus a constant, b

$$r_i = g_i(\text{FOE}) + h_i(\mathbf{P}) + b \quad (\text{A1})$$

or using vector notations

$$\mathbf{r} = \mathbf{g}(\text{FOE}) + \mathbf{h}(\mathbf{P}) + \mathbf{b} \quad (\text{A2})$$

where \mathbf{r} , \mathbf{g} , and \mathbf{h} are vectors with components r_i , $g_i(\text{FOE})$, and $h_i(\mathbf{P})$, and \mathbf{b} is a constant vector. In darkness, while the animal is pursuing, although there is no expansion pattern on the screen, the response vector takes the form

$$\mathbf{r}_{\text{dark}} = \mathbf{h}(\mathbf{P}) + \mathbf{c} \quad (\text{A3})$$

where \mathbf{c} is another constant vector.

Let us call $\mathbf{W}_{p^{\text{light}}}$ the weights of the optimal linear estimator of pursuit, $\hat{\mathbf{P}}_{\text{light}}$, obtained from the response of the neurons in conditions involving a combination of FOE and pursuit. Under the linear assumption, the conditional mean of $\hat{\mathbf{P}}_{\text{light}}$ is given by

$$\begin{aligned} E[\hat{\mathbf{P}}_{\text{light}}|\mathbf{P}] &= E[\mathbf{W}_{p^{\text{light}}}\mathbf{r}|\mathbf{P}] \\ &= \mathbf{W}_{p^{\text{light}}}E[\mathbf{g}(\text{FOE})|\mathbf{P}] + \mathbf{W}_{p^{\text{light}}}E[\mathbf{h}(\mathbf{P})|\mathbf{P}] + \mathbf{W}_{p^{\text{light}}}E[\mathbf{b}|\mathbf{P}] \end{aligned}$$

We now define $\hat{\mathbf{P}}_{\text{dark}}$ as the estimator of pursuit obtained by applying $\mathbf{W}_{p^{\text{light}}}$ to the activity in darkness, \mathbf{r}_{dark}

$$\begin{aligned} E[\hat{\mathbf{P}}_{\text{dark}}|\mathbf{P}] &= \mathbf{W}_{p^{\text{light}}}E[\mathbf{r}_{\text{dark}}|\mathbf{P}] \\ &= \mathbf{W}_{p^{\text{light}}}E[\mathbf{h}(\mathbf{P})|\mathbf{P}] + \mathbf{W}_{p^{\text{light}}}E[\mathbf{c}|\mathbf{P}] \end{aligned}$$

If we take the difference of the conditional means in the light and in the dark, we obtain

$$E[\hat{\mathbf{P}}_{\text{light}}|\mathbf{P}] - E[\hat{\mathbf{P}}_{\text{dark}}|\mathbf{P}] = \mathbf{W}_{p^{\text{light}}}E[\mathbf{g}(\text{FOE})|\mathbf{P}] + E[\mathbf{b}|\mathbf{P}] + E[\mathbf{c}|\mathbf{P}]$$

None of these terms depends on **P**.

Thus if FOE and pursuit interact linearly, this difference should be equal to a constant vector for all tested pursuit directions; that is, this difference should satisfy the above equation derived under the linear assumption. If they interact nonlinearly, the difference should be different for the different tested pursuit directions.

DISCLOSURES

W. Page and C. Duffy were supported by National Institutes of Health Grant R01EY-0287 and Training Grant T32EY-07125 to the Center for Visual Sciences, and S. Ben Hamed and A. Pouget were supported by MH-57823-06 and research grants from the Office of Naval Research and the McDonnell-Pew, Sloan, and Schmitt Foundations.

REFERENCES

- Baldi P and Heiligenberg W.** How sensory maps could enhance resolution through ordered arrangements of broadly tuned receivers. *Biol Cybern* 59: 313–318, 1988.
- Beintema JA and van den Berg AV.** Heading detection using motion templates and eye velocity gain fields. *Vision Res* 38: 2155–2179, 1998.
- Bradley DC, Maxwell M, Andersen RA, Banks MS, and Shenoy KV.** Mechanisms of heading perception in primate visual cortex. *Science* 273: 1544–1547, 1996.
- Bremmer F, Pouget A, and Hoffmann KP.** Eye position encoding in the macaque posterior parietal cortex. *Eur J Neurosci* 10: 153–160, 1998.
- Britten KH and Van Wezel RJ.** Area MST and heading perception in macaque monkeys. *Cereb Cortex* 12: 692–701, 2002.
- Deneve S, Latham PE, and Pouget A.** Reading population codes: a neural implementation of ideal observers. *Nat Neurosci* 2: 740–745, 1999.
- Deneve S, Latham PE, and Pouget A.** Efficient computation and cue integration with noisy population codes. *Nat Neurosci* 4: 826–831, 2001.
- Duffy CJ.** MST neurons respond to optic flow and translational movement. *J Neurophysiol* 80: 1816–1827, 1998.
- Duffy CJ and Wurtz RH.** Response of monkey MST neurons to optic flow stimuli with shifted centers of motion. *J Neurosci* 15: 5192–5208, 1995.
- Duffy CJ and Wurtz RH.** Planar directional contributions to optic flow responses in MST neurons. *J Neurophysiol* 77: 782–796, 1997.
- Edelman S and Poggio T.** Models of object recognition. *Curr Opin Neurobiol* 1: 270–273, 1991.
- Georgopoulos AP, Schwartz AB, and Kettner RE.** Neuronal population coding of movement direction. *Science* 233: 1416–1419, 1986.
- Gibson JJ.** *The Perception of the Visual World.* Boston, MA: Houghton Mifflin, 1950.
- Heeger DJ and Jepson AD.** Subspace methods for recovering rigid motion. I. Algorithm and implementation. *Inter J Comput Vision* 7: 95–117, 1992.
- Hildreth EC.** Recovering heading for visually-guided navigation. *Vision Res* 32: 1177–1192, 1992.
- Logothetis NK and Pauls J.** Psychophysical and physiological evidence for viewer-centered object representations in the primate. *Cereb Cortex* 3: 270–288, 1995.
- Morgan N and Boulard H.** Generalization and parameter estimation in feedforward nets: some experiments. In: *Advances in Neural Information Processing Systems 2*, edited by Touretzky DS. San Mateo, CA: Morgan Kaufmann, 1990, p. 630–637.
- Page WK and Duffy CJ.** MST neuronal responses to heading direction during pursuit eye movements. *J Neurophysiol* 81: 596–610, 1999.
- Page WK and Duffy CJ.** Heading representation in MST: sensory interactions and population encoding. *J Neurophysiol* 89: 1994–2013, 2003.
- Perrone JA and Stone LS.** Model of self-motion estimation within primate extrastriate visual cortex. *Vision Res* 34: 2917–2938, 1994.
- Poggio T.** A theory of how the brain might work. *Cold Spring Harbor Symp Quant Biol* 55: 899–910, 1990.
- Poggio T and Edelman S.** A network that learns to recognize three-dimensional objects. *Nature* 343: 263–266, 1990.
- Poggio T and Girosi F.** Regularization algorithms for learning that are equivalent to multilayer networks. *Science* 247: 978–982, 1990.
- Pouget A and Sejnowski TJ.** A neural model of the cortical representation of egocentric distance. *Cereb Cortex* 4: 314–329, 1994.
- Pouget A and Sejnowski TJ.** Spatial transformations in the parietal cortex using basis functions. *J Cogn Neurosci* 9: 222–237, 1997.
- Pouget A and Snyder LH.** Computational approaches to sensorimotor transformations. *Nat Neurosci* 3(Suppl): 1192–1198, 2000.
- Pouget A, Zhang K, Deneve S, and Latham PE.** Statistically efficient estimation using population coding. *Neural Comput* 10: 373–401, 1998.
- Rieger JH and Lawton DT.** Processing differential image motion. *J Opt Soc Am A* 2: 354–360, 1985.
- Royden CS.** Mathematical analysis of motion-opponent mechanisms used in the determination of heading and depth. *J Opt Soc Am A* 109: 2128–2143, 1997.
- Royden CS, Banks MS, and Crowell JA.** The perception of heading during eye movements. *Nature* 360: 583–585, 1992.
- Seung HS and Sompolinsky H.** Simple models for reading neuronal population codes. *Proc Natl Acad Sci USA* 90: 10749–10753, 1993.
- Shadlen MN, Britten KH, Newsome WT, and Movshon TA.** A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci* 16: 1486–1510, 1996.
- Siegel RM.** Representation of visual space in area 7a neurons using the center of mass equation. *J Comput Neurosci* 5: 365–381, 1998.
- Upadhyay UD, Page WK, and Duffy CJ.** MST responses to pursuit across optic flow with motion parallax. *J Neurophysiol* 84: 818–826, 2000.
- Warren WH Jr. and Hannon DJ.** Eye movements and optical flow. *J Opt Soc Am A* 7: 160–169, 1990.
- Zohary ESM and Newsome WT.** Correlated neuronal discharge rate and its implication for psychophysical performance. *Nature* 370: 140–143, 1994.