Single-Unit Activity in Cortical Area MST Associated With Disparity-Vergence Eye Movements: Evidence for Population Coding

A. TAKEMURA,[1] Y. INOUE,[1] K. KAWANO,[1] C. QUAIA,[2,3] AND F. A. MILES[2]
[1]*Neuroscience Section, Electrotechnical Laboratory, Ibaraki 305, Japan;* [2]*Laboratory of Sensorimotor Research, National Eye Institute, National Institutes of Health, Bethesda, Maryland 20892; and* [3]*Dipartimento di Elettronica, Elettrotecnica ed Informatica, Universitá degli Studi di Trieste, 34100 Trieste, Italy*

Received 3 August 2000; accepted in final form 3 January 2001

Takemura, A., Y. Inoue, K. Kawano, C. Quaia, and F. A. Miles. Single-unit activity in cortical area MST associated with disparity-vergence eye movements: evidence for population coding. *J Neurophysiol* 85: 2245–2266, 2001. Single-unit discharges were recorded in the medial superior temporal area (MST) of five behaving monkeys. Brief (230-ms) horizontal disparity steps were applied to large correlated or anticorrelated random-dot patterns (in which the dots had the same or opposite contrast, respectively, at the two eyes), eliciting vergence eye movements at short latencies [65.8 ± 4.5 (SD) ms]. Disparity tuning curves, describing the dependence of the initial vergence responses (measured over the period 50–110 ms after the step) on the magnitude of the steps, resembled the derivative of a Gaussian, the curves obtained with correlated and anticorrelated patterns having opposite sign. Cells with disparity-related activity were isolated using correlated stimuli, and disparity tuning curves describing the dependence of these initial neuronal responses (measured over the period of 40–100 ms) on the magnitude of the disparity step were constructed ($n = 102$ cells). Using objective criteria and the fuzzy c-means clustering algorithm, disparity tuning curves were sorted into four groups based on their shapes. A post hoc comparison indicated that these four groups had features in common with four of the classes of disparity-selective neurons in striate cortex, but three of the four groups appeared to be part of a continuum. Most of the data were obtained from two monkeys, and when the disparity tuning curves of all the individual neurons recorded from either monkey were summed together, they fitted the disparity tuning curve for that same animal's vergence responses remarkably well (r^2: 0.93, 0.98). Fifty-six of the neurons recorded from these two monkeys were also tested with anticorrelated patterns, and all showed significant modulation of their activity ($P < 0.005$, 1-way ANOVA). Further, when all of the disparity tuning curves obtained with these patterns from either monkey were summed together, they too fitted the disparity tuning curve for that same animal's vergence responses very well (r^2: 0.95, 0.96). Indeed, the summed activity even reproduced idiosyncratic differences in the vergence responses of the two monkeys. Based on these and other observations on the temporal coding of events, we hypothesize that the magnitude, direction, and time course of the initial vergence velocity responses associated with disparity steps applied to large patterns are all encoded in the summed activity of the disparity-sensitive cells in MST. Latency data suggest that this activity in MST occurs early enough to play an active role in the generation of vergence eye movements at short latencies.

INTRODUCTION

Vergence eye movements function to align both eyes on the same object and facilitate the binocular fusion of visual images. An important cue in this process is the slight difference in the locations of the two retinal images that arises from the slight difference in the viewpoints of the two eyes: binocular disparity. Most studies of disparity-induced vergence have examined the transfer of fixation between small targets presented at different distances and have reported latencies ranging from 150 to 200 ms in humans (Jones 1980; Mitchell 1970; Rashbass and Westheimer 1961; Westheimer and Mitchell 1956) and from 135 to 177 ms in monkeys (Cumming and Judge 1986). However, it has been shown that step changes in the horizontal disparity of a large pattern result in machine-like vergence responses with very short latencies: <60 ms in monkeys and <0.80 ms in humans (Busettini et al. 1996; Masson et al. 1997). It has been suggested that these disparity vergence responses are important for the rapid automatic correction of residual (i.e., small) vergence errors (Busettini et al. 1996). In line with this suggestion, these short-latency vergence responses have disparity tuning curves that resemble the derivative of a Gaussian, are well fit by odd (sine phase) Gabor functions and have a roughly linear servo region that extends only a degree or so on either side of zero disparity (Masson et al. 1997). It has recently been reported (Masson et al. 1997) that horizontal disparity steps applied to dense anticorrelated random-dot patterns (in which the patterns seen by the two eyes have opposite contrast so that each black dot in one eye is matched to a white dot in the other eye) also elicit short-latency vergence responses that are very similar to those observed with normal correlated patterns except that they are in the opposite direction. In line with the earlier observations of Cogan et al. (1993), Masson et al. also showed that all of their human subjects were able to discriminate between 1.2° crossed and uncrossed disparity steps applied to dense correlated patterns, but none was able to make these discriminations with dense anticorrelated patterns, which evoked strong binocular rivalry. Cumming and Parker (1997) have reported that monkeys too can make such discriminations with correlated patterns but not with anticorrelated patterns. These anticorrelated patterns

Address for reprint requests: A. Takemura, Neuroscience Section, Electrotechnical Laboratory, 1-1-4 Umezono, Tsukubashi, Ibaraki 305, Japan (E-mail: atakemur@etl.go.jp).

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.

therefore provide an interesting disparity stimulus in that they can support vergence eye movements but not depth perception.

It has often been suggested that disparity-induced vergence utilizes disparity-selective neurons to sense vergence errors, and such neurons have been recorded in various regions of the monkey cortex, including striate and extrastriate visual areas (Burkhalter and Van Essen 1986; Cumming and Parker 1999, 2000; Felleman and Van Essen 1987; Hubel and Livingstone 1987; Hubel and Wiesel 1970; Poggio and Fischer 1977; Poggio and Talbot 1981; Poggio et al. 1988; Prince et al. 2000; Smith et al. 1997; Trotter et al. 1996), as well as the middle temporal area (MT) (Bradley and Andersen 1998; Bradley et al. 1995; DeAngelis and Newsome 1999; DeAngelis et al. 1998; Maunsell and Van Essen 1983a), MST (Eifuku and Wurtz 1999; Roy et al. 1992), the posterior parietal area (Sakata et al. 1983), the lateral bank of the intraparietal sulcus (LIP) (Gnadt and Mays 1995), and the frontal eye fields (FEF) (Ferraina et al. 2000; Gamlin et al. 1996). Most of the earlier studies grouped cells according to the shapes of their disparity tuning curves using the classification scheme of Poggio and Fischer (1977), which recognized two general groupings of disparity-selective neurons, later termed "tuned" and "reciprocal" by Poggio et al. (1988). Tuned neurons had narrow tuning curves with either a peak ("tuned-excitatory" neurons) or a trough ("tuned-inhibitory" neurons) centered on the plane of fixation. Reciprocal neurons had asymmetric tuning curves and responded to a broad range of disparities in front ("near" neurons) or behind ("far" neurons) the plane of fixation. Subsequently, Poggio et al. (1988) recognized two additional tuned groups with peaks in front ("tuned near" neurons) or behind ("tuned far" neurons) the plane of fixation and this led to the renaming of the "tuned excitatory neurons" as "tuned zero neurons." Cumming and Parker (1997) recently showed that most disparity-selective neurons in cortical area V1 of monkeys also respond to the disparity of anticorrelated random-dot patterns, often with the opposite sign, e.g., neurons that had "tuned excitatory" disparity tuning curves with correlated patterns had "tuned inhibitory" tuning curves with anticorrelated patterns. Such responses are in line with the suggestion that these neurons act as purely local filters (Cumming and Parker 1997, 2000; Nomura et al. 1990; Ohzawa 1998; Ohzawa et al. 1990).

Our purpose in the present study was to see if there are neurons in area MST of the macaque monkey that have the necessary properties to initiate the short-latency disparity-vergence responses described by Busettini et al. (1996) and Masson et al. (1997). A strong motivating factor came from preliminary evidence that bilateral lesions in MST cause a significant reduction in these vergence responses (Takemura et al. 1999, 2000). Our study was restricted to the initial (open-loop) neuronal responses that occur before the associated vergence responses have had time to modify the central neuronal responses via the disparity feedback loop. We here report that MST neurons can be activated by horizontal disparity steps applied to either correlated or anticorrelated random-dot patterns at latencies that are probably short enough for these neurons to have a causal role in producing even the earliest vergence responses. However, comparatively few of these neurons had disparity tuning curves that resembled the tuning curves for vergence and, when sorted using a fuzzy clustering algorithm, the curves fell into four major groups, corresponding roughly to classes of disparity-selective neurons described in the visual cortex (Poggio et al. 1988). Thus qualitatively at least, most individual tuning curves resembled those at earlier, overtly sensory, stages. Interestingly, when these curves were simply summed together, they fitted the tuning curves for the vergence responses elicited by both correlated and anticorrelated stimuli, indicating that the associated motor responses are encoded at the population level. Nonetheless, using a genetic algorithm, it was possible to identify subsets of neurons whose tuning curves, when summed together, gave an even better fit to the vergence tuning curves, though these subsets invariably included cells from all four groups. Additional analyses of the spike trains elicited by disparity steps revealed considerable variation across cells in the latency, amplitude, and time course of the changes in discharge rate. When all of the spike trains elicited by a given disparity step were summed together to give an average discharge profile for the whole population of cells, many were rather noisy, but others that were less so matched the temporal profile of the motor response, vergence velocity, quite well. Based on these findings, we hypothesize that the disparity-sensitive cells in MST each encode only some aspect(s) of the sensory input and/or motor output, but that the population of cells as a whole encodes the complete motor output (vergence velocity).

Preliminary results have been presented elsewhere (Takemura et al. 1997, 1998, 1999).

METHODS

We recorded single unit activity in the MST area of five adolescent Japanese monkeys (*Macaca fuscata*, 6–8 kg) in response to disparity steps applied to large random-dot patterns. Prior to any surgery, animals were trained to fixate small target spots on a tangent screen for a liquid reward using the dimming task of Wurtz (1969). After the completion of training, animals were anesthetized with pentobarbital sodium for the surgical implantation (under aseptic conditions) of *1*) a pedestal, secured to the skull to permit the head to be fixed in the standard stereotaxic position during the experiments, *2*) a cylinder, secured to the skull over the superior temporal sulcus for the chronic recording of single neuron activity, and *3*) scleral search coils, around both eyes for recording eye movements (Judge et al. 1980). All experimental procedures were approved by the Electrotechnical Laboratory Animal Care and Use Committee and have been fully described elsewhere (Kawano et al. 1994).

Disparity stimuli

During recording sessions, which were each several hours long, the animal sat in a primate chair with the head secured and faced a translucent tangent screen onto which two identical random-dot patterns were back-projected. The screen was 50 cm in front of the eyes and subtended 90° along the vertical and horizontal meridia. Orthogonal polarizing filters in the two projection paths and in front of each eye ensured that each pattern was visible to only one eye. Mirror galvanometers in the two light paths were used to control the horizontal positions of the two images (binocular disparity). Using fluid reinforcement, animals were rewarded for fixating stationary red target spots, which were projected onto the patterns on the screen. Binocular disparity stimuli were applied to random-dot patterns that were either exactly matched at the two eyes (standard correlated patterns) or of opposite contrast (anticorrelated patterns).

STANDARD PARADIGM USING CORRELATED PATTERNS. At the beginning of each trial, the patterns seen by the two eyes were identical, overlapped exactly (zero binocular disparity), and filled the

screen. The patterns consisted of white dots on a black background (dots had a luminance of 0.8 cd/m^2, a diameter of 1.5° and covered 50% of the image space). Horizontal disparity steps (crossed and uncrossed, ranging in amplitude from 0.5 to 6.0°) were applied by displacing the two images equally in opposite directions. Because previous studies had shown that the vergence responses were subject to transient postsaccadic enhancement (Busettini et al. 1996), these steps were applied 50 ms after 10° leftward centering saccades guided by target spots projected onto the screen. The experimental situation was the same as that used by Busettini et al. (1996) in all essentials. Because we were interested only in the initial vergence responses, the disparity steps lasted only 230 ms, and, if there were no saccades during this time, then the data were stored on a hard disk and the animal was given a drop of water; otherwise, the trial was aborted and fluid was withheld. At this point, both images were blanked for 500 ms by mechanical shutters and then reappeared once more for the start of the next trial. Note that all experiments included control trials in which no steps were applied (saccade-only trials). By applying the disparity steps in the immediate wake of centering saccades, we ensured that the animal was alert during the steps, the stimulus pattern was always centered on the retina at the onset of the steps, and the vergence responses were subject to postsaccadic enhancement.

PARADIGM USING ANTICORRELATED PATTERNS. For this paradigm, the right eye saw the same white dots on a black background as when the binocular patterns were correlated, and the left eye saw a matching negative image (black dots on a white background). However, trials started with the screen a featureless gray (with the same space-averaged luminance as for the patterns), until 50 ms after a 10° leftward centering saccade (again, guided by projected target spots), at which time the anticorrelated random-dot patterns suddenly appeared with a fixed horizontal disparity. Here too the disparity stimuli were presented only briefly (230 ms) before the screen was blanked, ending the trial. This procedure for applying disparate anticorrelated stimuli was the same as that used by Masson et al. (1997) and was necessitated because disparity steps applied directly to anticorrelated patterns at best elicit only weak vergence eye movements (unpublished observations).

Data collection

The techniques for recording unit activity in MST were the same as previously described (Kawano et al. 1992, 1994) and will only be given in brief here. A hydraulic microdrive (Narishige Mo-9) was mounted on the recording cylinder, and glass-coated tungsten microelectrodes were used for the initial identification and mapping of the MT/MST region. Subsequently, a fixed grid system (Crist et al. 1988) was used to position a guide tube through which a flexible tungsten microelectrode was introduced into the MST area for single-cell recordings. The tip of the guide tube was positioned 3–5 mm above the MST. Neuronal activity was recorded using standard extracellular techniques. Spikes were detected with a time-amplitude window discriminator with a resolution of 1 ms. We selectively isolated neurons whose discharge was modulated by disparity steps applied to correlated patterns, and only after obtaining ≥40 samples of the responses to the complete set of disparity steps did we record responses to similar "steps" applied to anticorrelated patterns. Eye velocity signals were sampled at 500 Hz and all other analog signals at 250 Hz. All data were transferred to a work station (SunSparc) for quantitative analysis.

Data analysis

Horizontal vergence was computed by subtracting the horizontal position signal for the right eye from that of the left eye. We used the convention that rightward positions and velocities are positive, hence, crossed disparities and convergence were also positive. The latency of the vergence eye movements (and associated neuronal responses) was taken to be the time when the mean vergence acceleration (and the mean discharge rate) first exceeded the baseline noise by 2 SD. Vergence responses were quantified by measuring the change in vergence position over the 60-ms time period beginning 50 ms after stimulus onset, and disparity tuning curves were constructed by plotting these measures against the amplitude of the disparity step. To quantify the neuronal responses, we first constructed peristimulus time histograms (binwidth, 1 ms) from the responses to multiple presentations of each disparity step, computing a spike density function by convolving each spike with a Gaussian pulse whose standard sigma was 3 ms (MacPherson and Aldridge 1979; Richmond et al. 1987). These histograms were then used to compute the (average) instantaneous discharge frequency over time. The mean discharge frequency over the 60-ms period starting 40 ms after stimulus onset was computed for the responses to each disparity step, and these values were then plotted against the amplitude of the step: these plots will be referred to as disparity tuning curves. All data shown in the figures, including both eye movements and neuronal discharges, have had the responses on saccade-only trials (i.e., no disparity step applied) subtracted to eliminate any postsaccadic vergence drift and postsaccadic neuronal activity. This has the effect of forcing all the disparity tuning curves through the origin.

FUZZY CLUSTERING ANALYSIS. It has been usual to group disparity-selective neurons according to the shape of their disparity tuning curve using a classification scheme described by Poggio and coworkers (Poggio and Fischer 1977; Poggio et al. 1988). This scheme has been very successful but relies on a subjective assessment. In an effort to classify the cells according to the shape of their tuning curve in an objective manner, we turned to clustering methods. These methods attempt to achieve an optimal partitioning of the data into groups. When visual inspection of the data reveals the presence of relatively tight and separate clusters, classic clustering algorithms successfully assign each datum to one of several discrete groups. But when a human observer is uncertain about the separation between groups, as was the case with our data set, these algorithms may not find any structure in the data. In this case, one can try fuzzy clustering algorithms, which retain much more information about the distribution of the data being clustered. We implemented the fuzzy c-means clustering algorithm of Bezdek (1981), and the technical details are given in APPENDIX A. Here, we provide only a brief outline of the general operations performed.

As we were interested in grouping the cells' responses solely on the basis of the shapes of their disparity tuning curves, we first normalized the individual curves by adjusting the gain and bias so that all had the same peak-to-peak amplitude and mean. It was then necessary to decide how many groupings would be used by specifying the number of clusters, *n*, and the rationale for the number of clusters finally used (4) is given in RESULTS. The algorithm begins by randomly selecting *n* "center" waveforms from within the available data space, one for each cluster, and then computes *n* "membership values" for each of the tuning curves, one for each of the *n* clusters. These membership values are a measure of the "similarity" of a given curve to each of the *n* "center" curves and are constrained such that each membership value can range from 0 to 1, the more similar the curve to the center curve of a given cluster the higher its membership in that cluster, and the sum of the *n* membership values for each tuning curve is 1. Through an iterative process, the algorithm selects cluster centers to maximize the homogeneity within each cluster and the heterogeneity between clusters. The algorithm generally converged after 10–20 iterations, and we then assigned each cell to one of *n* groups based on the cluster in which the cell had its highest membership: cells whose membership values peaked in *cluster 1* were placed in *group 1*, cells whose membership values peaked in *cluster 2* were placed in *group 2*, and so forth.

GENETIC ALGORITHM. In examining the correlation between the single-unit activity and the vergence eye movements elicited by

disparity steps, we attempted to identify the subset of cells whose disparity tuning curves when summed together best matched the disparity tuning curve for vergence. However, if the number of units recorded in a given animal is n, then 2^n subsets are possible. In the present study, for example, we obtained disparity tuning curves for 49 units from one animal, and we estimate that our current computers would require hundreds of years to examine all 2^{49} subsets. After reviewing the options, we decided to use a genetic algorithm (GA). The rationale behind this choice is that for problems in which there is a large number of binary variables (as here), GAs are known to outperform all other methods (computation time being equal): unlike other optimization algorithms, GAs sample several regions of the solution space in parallel. The technical details are given in APPENDIX B and only a brief outline of the general operations performed is provided here.

The data from each animal were treated separately because of small differences in the shapes of their disparity tuning curves for vergence. In our implementation of the GA algorithm, each "chromosome" in effect represents one of the 2^n subsets of tuning curves. We started with 5,000 chromosomes, each having the same complement of n "genes," one for each of the n units recorded in the animal under scrutiny. All genes were randomly assigned a value of either 1, indicating that they made a contribution ("were expressed"), or 0, indicating that they made no contribution ("were not expressed"). For each chromosome, the disparity tuning curves of the units/genes that had been assigned a value of 1 were summed together and then fitted to the disparity tuning curve for the vergence response of that monkey. The task of the GA was to then "evolve" a chromosome (or chromosomes) whose subset of "expressed" tuning curves when summed together best fit the vergence data. The mean squared error (MSE) was used to assess the goodness of fit, and a histogram showing the distribution of MSEs among the first generation of 5,000 chromosomes invariably indicated a wide range of values. New generations of chromosomes (each having 5,000 chromosomes) with progressively smaller MSEs were created using a set of standard evolutionary rules. The algorithm ran for 50 generations, during which the minimum MSE for the population gradually diminished, usually stabilizing after ~30 generations. When the last (50th) generation was reached, the vast majority of the "chromosomes" had the same string of "expressed genes," and further evolution was virtually impossible. This surviving string of expressed genes represents the algorithm's estimate of the subset of units whose summed disparity tuning curves best correlate with the disparity tuning curve for vergence.

Histology

At the conclusion of recordings in a given monkey, that animal was deeply anesthetized with pentobarbital and perfused through the heart with saline followed by 10% Formalin. The animal's brain was removed, and frozen sections were cut at 50 μm in the sagittal plane, mounted on microscope slides, and stained with cresyl violet for cell bodies and with a modified silver stain (Gallyas 1979) for myelinated fibers. Electrolytic lesions facilitated histological reconstruction of the electrode tracks, and recording sites were verified using both Nissl and myelination. Sample electrode tracks were verified to pass through the MST using X-rays, and neurons were assumed to be in the MST based on their physiological characteristics (such as preferred speed and receptive field size).

RESULTS

Initial vergence responses to disparity steps (correlated patterns)

Horizontal disparity steps applied to large-field, correlated, random-dot patterns elicited vergence eye movements at short latency that were, in all essentials, like those described by Busettini et al. (1996) and Masson et al. (1997). Latency estimates for the five monkeys, based on the responses to the optimal disparity steps, were 73 ms (*monkey H,* +2°), 68 ms (*monkey L,* +2°), 65 ms (*monkey M,* −2°), 58 ms (*monkey N,* +1°), and 59 ms (*monkey Q,* +1°).[1] Although no attempt was made to obtain formal estimates of latencies to nonoptimal stimuli, the vergence velocity temporal profiles suggested that for a given animal, latency was largely independent of the stimulus amplitude: see, for example, the sample profiles from *monkey H* in Fig. 1, those in *A* showing responses to crossed disparity steps, and those in *B* responses to uncrossed disparity steps, the stimuli ranging in amplitude from 0.5 to 6°. The profiles in Fig. 1*A* reach a peak and then decline, often before the closing of the disparity feedback loop (twice the response latency), whereas the profiles in Fig. 1*B* have a more varied time course and some fail to reach a peak within the time window shown (150 ms after the disparity step). The initial responses to small (<2°) steps were always compensatory in that they operated to reduce the seen disparity: small crossed steps elicited increases in the vergence angle and small uncrossed steps the converse. Responses were generally maximal with steps of 1–3° and declined with larger steps, sometimes showing reversal (Fig. 1*B*). These trends are evident from the disparity tuning curves in Fig. 1*C*, in which the change in vergence position (measured over the time period of 50–110 ms) is plotted against the amplitude of the disparity step. These plots indicate that the system showed appropriate servo-like behavior only for small disparity steps. That is, increases in the input resulted in roughly proportional increases in the output (in the compensatory direction) only for steps of less than a degree or so. (Note that the amplitude and direction of the responses of any given monkey to the largest steps were generally independent of whether the steps were crossed or uncrossed and showed considerable variation from one animal to another.) The disparity tuning curves resembled the derivative of a Gaussian and were well fit by a Gabor function: the parameters for the least squares best fits are listed for all five monkeys in Table 1. To evaluate the goodness of the fit, we computed the fraction of the disparity-induced variation in the data accounted for by the fit, r^2, as previously done by others (e.g., Cumming and Parker 1999). We found that r^2 ranged from 0.88 to 0.99, indicating that 88–99% of the variation in the data were captured by the fit. (Note that the Gabor functions for *monkeys N* and *Q* are plotted as the continuous lines in Fig. 6, *A* and *B*.)

Initial neuronal responses to disparity steps (correlated patterns)

We recorded the activity of 586 MST neurons in seven hemispheres of five monkeys while horizontal disparity steps were applied to correlated large-field random-dot patterns. About 20% of the neurons (122) responded to disparity steps at short latency, and the changes in discharge rate ranged from very transient—some lasting little more than 20 ms—to tonic. This is evident from the sample responses in Fig. 2, which shows the discharge rate temporal profiles of 20 units recorded from *monkey N* in response to 2° crossed-disparity steps. If a neuron has a response latency of L_n ms, and the associated vergence has a latency of L_v ms, the earliest time at which the

[1] A lower luminance level (0.5 vs. 0.8 cd/m^2) was used for the first three monkeys (*H, L, M*) and might account for the longer latencies in these animals.

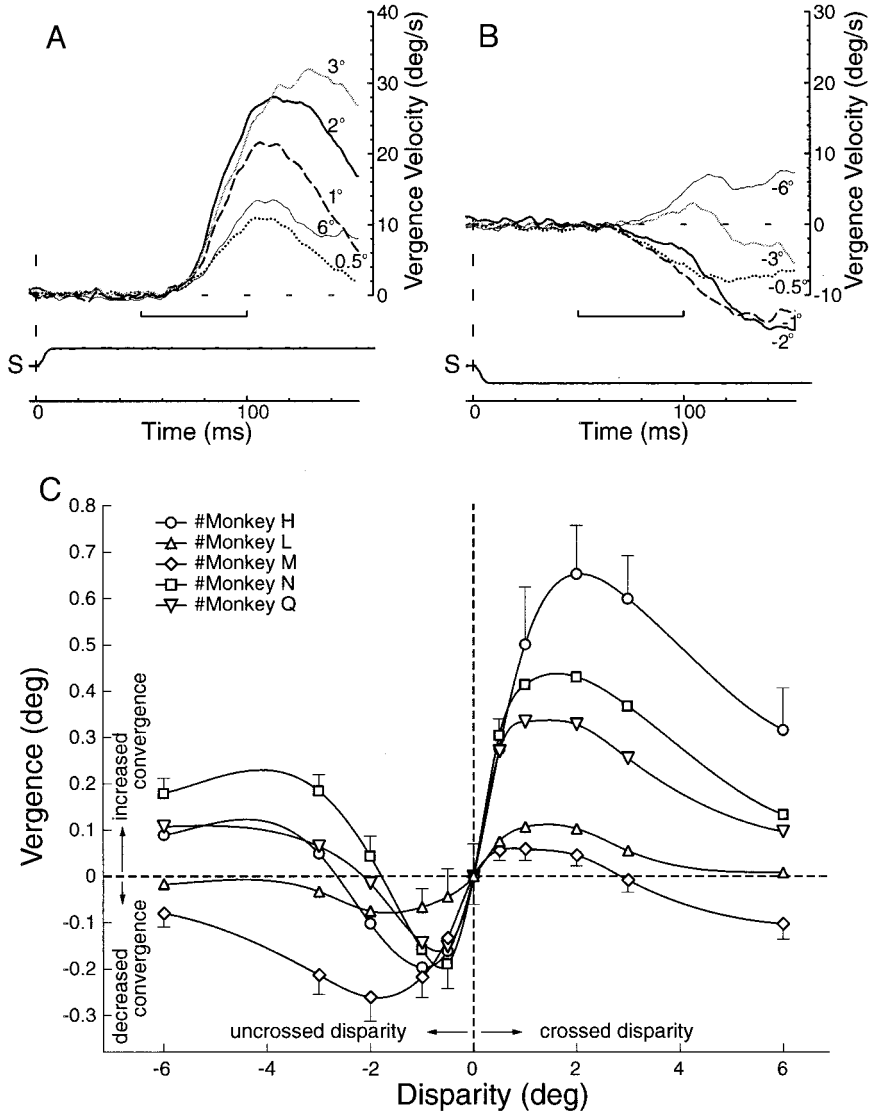

FIG. 1. Vergence responses: dependence on the amplitude and direction of the disparity step (correlated stimuli). *A*: traces show vergence velocity over time in response to crossed disparity steps. *B*: traces show the same for uncrossed steps. Upward deflections denote increased vergence and the numbers on the traces indicate the magnitudes of the steps in degrees. *C*: the mean change in vergence position in degrees, measured over the 60-ms period starting 50 ms after the disparity step (see the horizontal bars in *A* and *B*), is plotted against the magnitude of the step in degrees, for each of the 5 monkeys: disparity tuning curves. Continuous lines are spline interpolations with values at +6° disparity replicated at disparities of +9 and +12°, and values at −6° disparity replicated at disparities of −9 and −12°, and external knots at values with disparities of −18, −15, +15, and +18°. This method was preferred to classic spline interpolation (or cubic interpolation) because it dampens oscillations between absolute disparities of 3 and 6° (where there were no data). Error bars, 1 SD.

neuronal response could be affected by the decrease in retinal disparity secondary to the vergence response would be $L_n + L_v$ ms (after the disparity step). The estimated latency of the vergence response (see METHODS) in Fig. 2 was 59 ms (indicated by the dashed vertical line), which means that only neurons with a latency of <51 ms could have been affected by the closing of the disparity-feedback loop within the time window shown (110 ms after the disparity step). Seven of the

TABLE 1. *Parameters of the best-fit Gabor functions*

Monkey	Stimuli	Data	G	θ	k	n	x_{off}	y_{off}	r^2
H	Corr	Unit (11)	0.48	−134	2.45	0.10	0	0.12	0.94
		Verg	0.63	−107	2.06	0.10	0	0.07	0.95
L	Corr	Unit (8)	0.43	−99	1.70	0.04	0	−0.05	0.97
		Verg	2.99	−91	1.65	0.007	0	−0.08	0.96
M	Corr	Unit (3)	0.27	−73	1.80	0.11	0	−0.10	0.88
		Verg	0.24	−75	2.42	0.11	0	−0.09	0.97
N	Corr	Unit (49)	0.45	−106	1.24	0.13	0	0.03	0.95
		Verg	0.40	−107	1.51	0.13	0	0.03	0.97
	Anticorr	Unit (25)	0.42	72	0.93	0.11	−0.14	−0.18	0.99
		Verg	0.41	70	0.95	0.10	−0.06	−0.18	0.99
Q	Corr	Unit (31)	0.38	−111	1.50	0.15	0	−0.004	0.94
		Verg	0.46	−100	1.49	0.12	0	−0.014	0.96
	Anticorr	Unit (31)	0.20	59	1.45	0.16	−0.91	−0.36	0.96
		Verg	0.42	69	1.00	0.08	−0.90	−0.38	0.98

The Gabor function ($y = Ge^{-[(x-x_{off})^2/2k^2]} \cos [2\pi n(x - x_{off}) + (2\pi\theta/360)] + y_{off}$) was fitted to the disparity tuning curves for the summed activity (Unit) and for the vergence responses (Verg) obtained with correlated (Corr) and anticorrelated (Anticorr) stimuli. Number of units contributing to the summed activity are in parentheses. Some of these functions are plotted in Figs. 6 and 12.

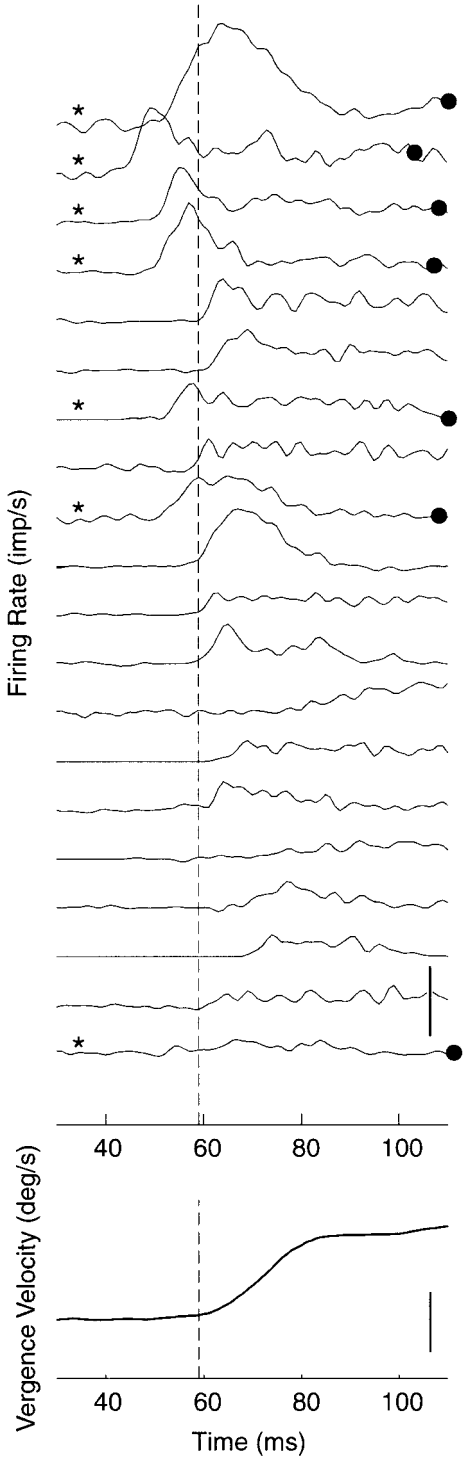

FIG. 2. Time course of the neuronal responses (correlated stimuli). *Top*: the changes in mean discharge rate over time in response to 2° crossed disparity steps for each of the 20 units that modulated most with this stimulus (ranked in descending order of their mean discharge rates over the period 30–110 ms after the step). *Bottom*: the changes in vergence velocity over time; the vertical dashed line shows the estimated latency of the vergence response (59 ms). Seven of the units (*) had response latencies <51 ms; ●, the estimated times at which the closing of the disparity feedback loop could first influence the discharges of these units (see text). The response latencies of the remaining units were too long for the disparity-feedback loop to close during the 110 ms time window shown. Calibration bars: 500 imps/s, 5°/s. Data from *monkey N*.

neurons in Fig. 2 had a short enough latency to have been so affected (see *), but even in those cases, the loop wouldn't have closed until long after the neurons' initial burst of activity had ended. This is apparent from the estimated time of closure of the disparity feedback loop for those seven cells (indicated by ● on the relevant traces in Fig. 2). A similar scrutiny of the entire population of cells indicated that the discharges of ~60% had a phasic component, which in every case was independent of the closure of the disparity feedback loop. The histograms in Fig. 3 show the distributions of the latency estimates for the neurons that met the response-onset criterion (2 SD above the mean control level: see METHODS). The estimates in Fig. 3*A* are given with respect to the onset of the disparity steps ($n = 71$; median, 55 ms; range 43–86 ms), whereas those in Fig. 3*B* are given with respect to the (measured) onset of the vergence responses. The onset of the neuronal responses preceded the onset of the vergence eye movements in 43/50 units (86%) by ≤24 ms (median lead, 9 ms). Note that the measures in Fig. 3 were all obtained from the

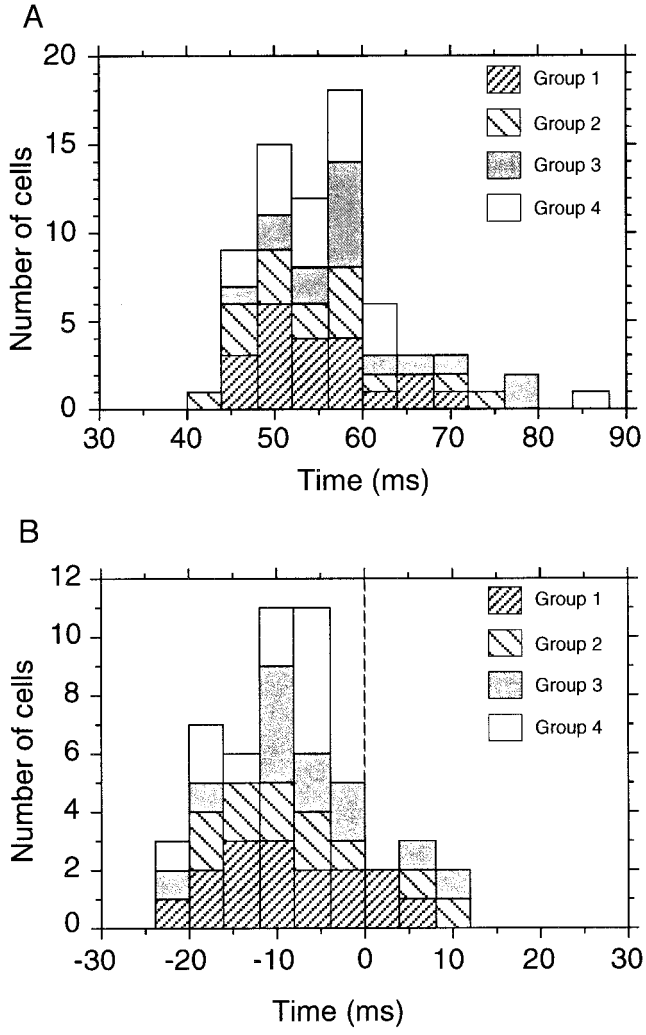

FIG. 3. Latency of onset of discharges (correlated stimuli). Histograms show the distribution of latencies with respect to the onset of the stimulus (*A*) and the onset of the associated vergence eye movements (*B*). For the latter, negative times indicate that the discharge preceded the vergence response.

response profiles elicited by the stimulus that was optimal for the cell (see METHODS).

DISPARITY TUNING CURVES OF SINGLE CELLS. Neuronal responses were quantified by measuring the mean discharge frequency over the 60-ms period starting 40 ms after the disparity step. Of the 122 MST neurons that showed sensitivity to horizontal disparity steps at short latency, 102 yielded sufficient data to allow complete disparity tuning curves to be plotted. These curves showed a variety of forms and, after normalization, we used the fuzzy c-means algorithm to organize them into groups based on their shapes (see METHODS). This algorithm sorted the curves into four clusters, and we then assigned each curve to one of four groups based on the cluster in which each curve had its largest membership. We chose to partition the curves into four groups because this was the largest number to yield a consistent grouping when the algorithm was run multiple times ($\geq$50).

Figure 4 shows the normalized disparity tuning curves for all 102 units arranged in the four groups, together with the average disparity tuning curves for the vergence eye movements of the two monkeys (*N* and *Q*) that yielded most (80/102, 78%) of the data (*bottom*). The general shapes of the tuning curves within the four groups conform very roughly to four of the six classes of disparity selective units described by Poggio et al. (1988): *1*) Cells in *group 1* tend to show increased activity in response to a somewhat restricted range of uncrossed disparities, their tuning curves having relatively narrow peaks that are distributed mostly between 0 and $-1°$ disparity and skewed to the left (toward uncrossed disparities), c.f., the "tuned far" cells of Poggio et al. (1988). *2*) Cells in *group 2* tend to show increased activity in response to crossed or uncrossed disparities, their tuning curves having a trough centered close to 0° disparity, cf., "tuned inhibitory" cells (though in the present case the trough is not the result of inhibition). *3*) Cells in *group 3* tend to show increased activity in response to a rather wide range of crossed disparities, their tuning curves having relatively broad peaks that are distributed between $+1$ and $+3°$ disparity and skewed to the right (toward crossed disparities), c.f., "near" cells. *4*) Cells in *group 4* tend to show increased activity in response to a somewhat restricted range of crossed disparities, their tuning curves having relatively narrow peaks that are distributed between 0 and $+1°$ disparity and skewed to the right (toward crossed disparities), cf., "tuned near" cells.

Although useful to distinguish between the groups, simple descriptors like near, tuned far, and so forth do not capture the full complexity of the shapes of the tuning curves within each group. For example, cells in the near *group 3* often show some increase in discharge with large uncrossed steps, and cells in the tuned far *group 1* often show some increase in discharge with large crossed steps.

Discrete groups or a continuum? The extent to which the curves within a given group have significant membership in

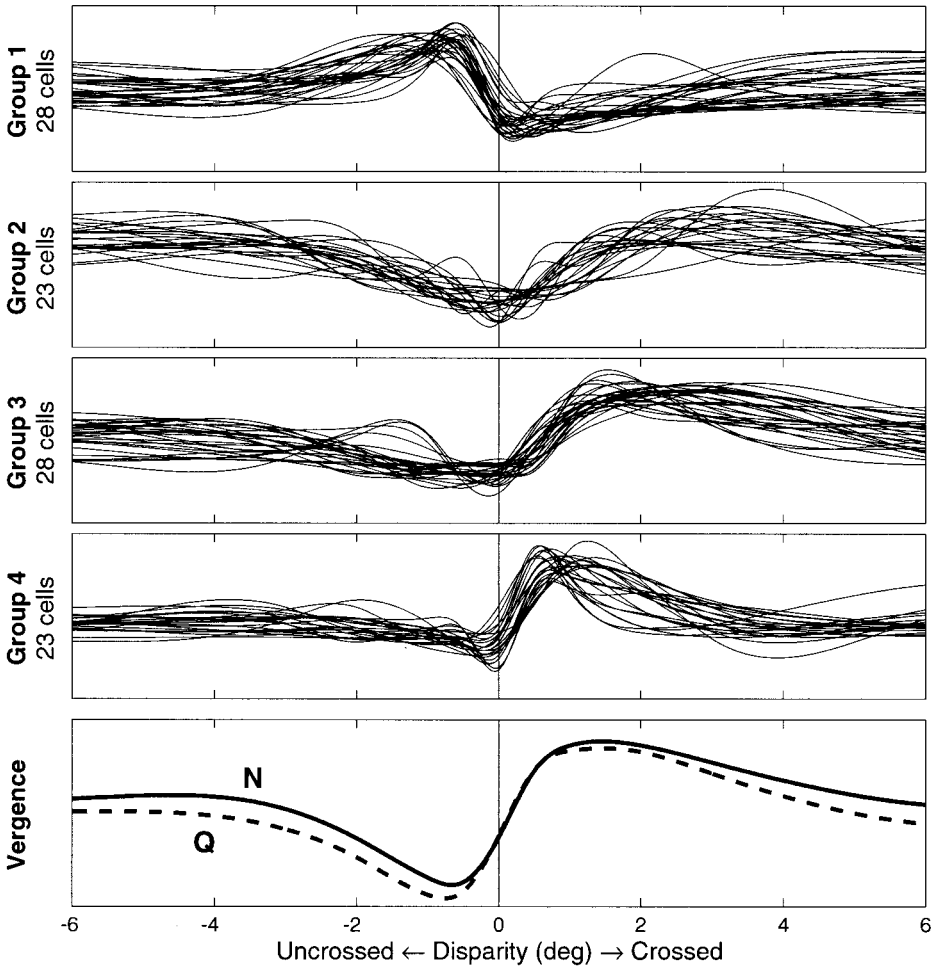

FIG. 4. Disparity tuning curves for the individual cells (correlated stimuli). *Top 4 graphs*: mean change in discharge rate over the 60-ms period starting 40 ms after the disparity step is plotted against the magnitude of the disparity step; *curves* are normalized and arranged in 4 groups based on the outcome of the fuzzy c-means clustering algorithm. *Bottom*: the disparity tuning curves for the vergence responses of the two monkeys that yielded most of the data (*N* and *Q*). Traces are spline interpolations (see legend of Fig. 1 for details).

only the "defining" cluster (that is, in *cluster 1* for *group 1*, in *cluster 2* for *group 2*, and so forth) provides some measure of the extent to which that group constitutes a discrete entity. By this token, the tuned far *group 1*, with an average membership in *cluster 1* of 0.75, was the most discrete and the near *group 3*, with an average membership in *cluster 3* of only 0.57, was the least discrete. All of the groups are to some degree fuzzy, however, insofar as all have cells with significant membership in more than one cluster. (Note that a cell would be maximally fuzzy if all of its membership values were 0.25.) This leaves open the possibility that we are dealing with a continuum, and there is some coarse hint of this in Fig. 4. Thus in passing from one group to another in the sequence, tuned near—near—tuned inhibitory—tuned far, some gradual trends are evident: the responses to crossed steps start with a peak at small disparities, then the peak flattens and shifts to higher disparities before largely disappearing, while the responses to uncrossed disparities show a similar trend but in the reverse order. If we are dealing with a "smooth" continuum then these trends should also be apparent within the groups.

Figure 5 (*top*) plots the memberships in each of the four clusters for all 102 cells. The cells are subdivided into the four groups and, within each group, are ranked according to their memberships in the defining cluster of an adjacent group.[2] Figure 5 (*bottom*) shows all of the disparity tuning curves in a color-coded form arranged along the abscissa in the order in which they are plotted in Fig. 5, *top*, with the disparity step represented along the ordinate (crossed steps upward); increases in activity are shown in red, zero activity in white, and decreases in activity in blue. The general impression is that the tuned far *group 1* is distinct, with a sharp discontinuity at the boundary with *group 2* (see particularly the responses to uncrossed disparities ranging from 0 to −2°) whereas the other three groups lie along a continuum.

We further examined the discreteness of the four groups using a hierarchical clustering algorithm. This algorithm starts out with as many clusters as there are data points and then progressively reduces the number by combining clusters that fail to achieve a set criterion for separation. This iterative process stops when the separation between the remaining clusters exceeds the criterion level. When run on our data, this algorithm did identify four clusters that were essentially identical to the four groups that we identified using the fuzzy c-means clustering algorithm. However, the algorithm did not stop there but went on to first combine *clusters 3* and *4* and then to combine this new cluster with *cluster 2* before reaching the criterion level for separation. Thus this algorithm reinforced the view that *group 1* is separate from all others and that *groups 2–4* have considerable overlap. Even though this provides some support for the notion that *groups 2–4* represent a continuum, we feel that, for descriptive purposes, it is still useful to treat them separately: whether they are samples from a continuum or from discrete entities, the three groups show clear differences, especially between the extremes (*groups 2* and *4*). Note also that there were no major differences in the latency distributions among the four groups of neurons: see the different patterns of shading in Fig. 3.

DISPARITY TUNING CURVES FOR (SUMMED) POPULATIONS OF CELLS. A characteristic feature of the disparity tuning curves for vergence in Fig. 1*C* is that they are roughly odd functions, that is, $f(x) = \approx -f(-x)$, with a linear segment passing smoothly through zero disparity defining the critical servo range over which changes in the input (binocular disparity) elicit roughly proportional changes in the output (vergence eye movement). It is significant that very few cells had disparity tuning curves with this exact same feature: The tuned inhibitory *group 2* cells have rather flat curves in the vicinity of zero, conveying little information about disparity in this range, and the tuning curves for most other cells undergo a substantial change in slope near zero disparity. Exceptions are small numbers of cells in the tuned near *group 4* and tuned far *group 1*, with positive and negative slopes, respectively, extending on either side of zero disparity. (Given appropriate excitatory or inhibitory connections, neurons with either positive or negative slopes could make a positive contribution to the vergence responses. That the tuning curves for vergence had a positive slope around zero disparity was determined solely by our sign convention.) To obtain an estimate of how well the vergence responses were encoded in the discharges elicited in individual MST cells, we fitted the disparity tuning curves of the individual cells to the tuning curves for vergence, the only free parameters being gain and *y* offset. We assumed that all neurons made a positive contribution to the vergence responses elicited by disparity steps in the important servo range, ±1° and so reversed the sign of those curves that had negative slopes over that range: this involved 29/102 cells, including all 28 in the tuned-far *group 1* and 1 in the tuned inhibitory *group 2*. The least-squares best fits ranged widely (r^2: mean, 0.51; range, 0.009–0.92), but the very best of the fits were good enough to raise the possibility that some cells might be more properly regarded as "vergence-related" rather than "disparity-related." We will revisit this issue later when we describe the responses of these cells to anticorrelated stimuli.

Population coding? To see how well the vergence responses were encoded in the discharges of the entire population of MST cells, we summed the raw (that is, nonnormalized) tuning curves for all units and then determined how well this population average fitted the tuning curve for vergence when gain and *y* offset were once more the only free parameters (so that the contributions of all neurons were given equal weight). Because the shapes of the disparity tuning curves for vergence differed slightly from one animal to another, it was necessary to treat the data from each animal separately, and we had adequate numbers of units to permit this for only 2/5 animals: *monkey N* (49 units) and *monkey Q* (31 units). We again reversed the sign of those curves that had negative slopes over the disparity range ±1°; this involved 17/49 cells in *monkey N* and 7/31 cells in *monkey Q*, all in the tuned far *group 1*. Given our sign convention (convergence is positive, divergence is negative), this meant that cells with positive slopes (e.g., tuned near) contributed to convergence and cells with negative slopes (tuned far) contributed to divergence: push-pull configuration. The resulting summed activity showed a very good fit to the

[2] The membership values are unsigned scalar entities, indicating only the similarity/proximity of the curves to each of the four center curves (see METHODS), and hence the cells within a group cannot be ranked by their memberships in their own defining cluster. However, they can be ranked by their memberships in any other, nondefining cluster because the curves within a given group must all lie on the same side of the 12-dimensional plane that passes through the center of the nondefining cluster and that is orthogonal to the line that goes through the center of the two clusters.

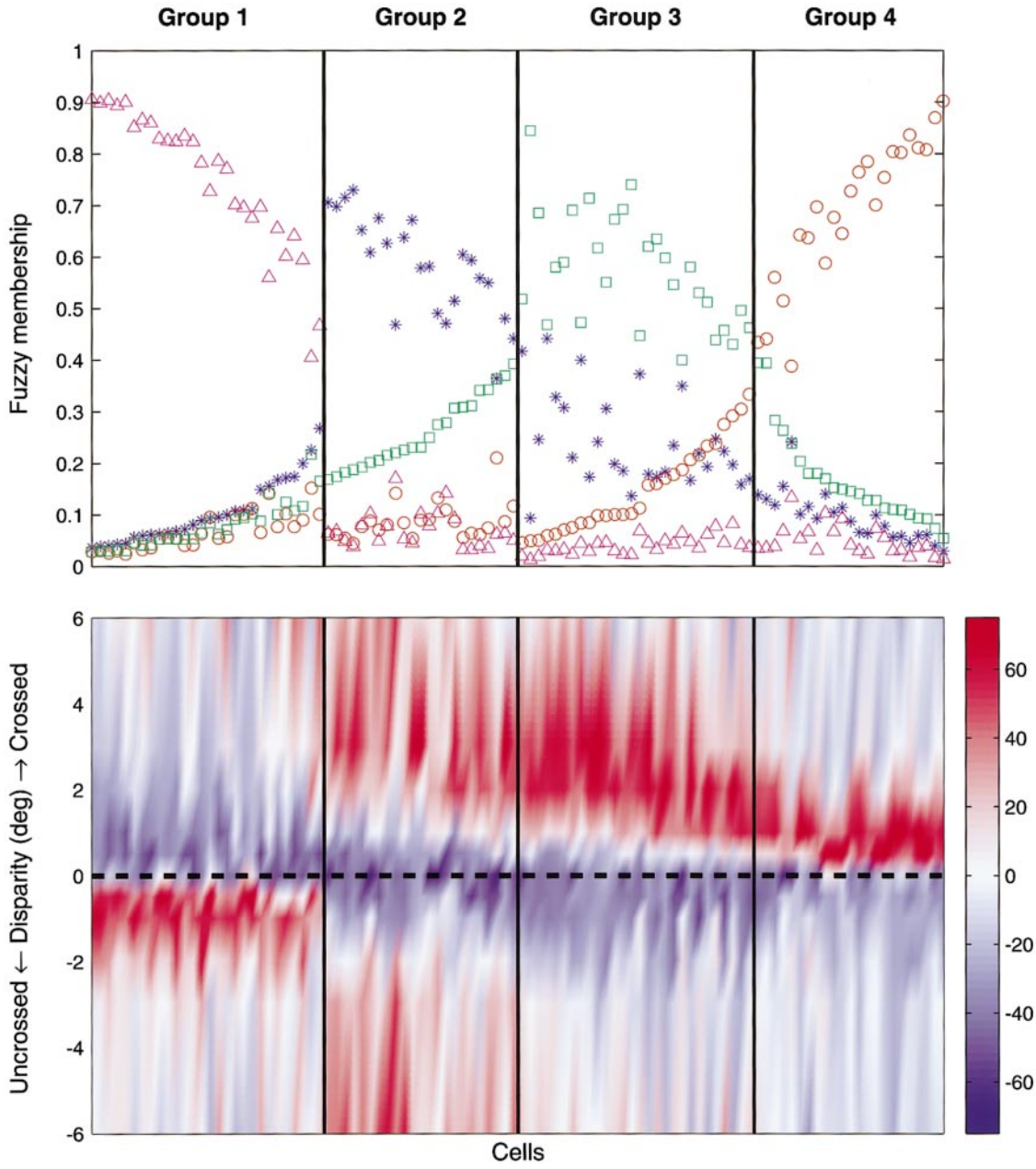

FIG. 5. Rank ordering of disparity tuning curves based on their membership values from the fuzzy clustering analysis (correlated stimuli). *Top*: membership values; each *point* along the *abscissa* represents an individual unit, whose four membership values (1 for each of the 4 clusters) are plotted along the *ordinate*; each unit is placed in the group that corresponds to the cluster in which it has its highest membership (*group 1* have their highest memberships in *cluster 1*, *group 2* their highest in in *cluster 2*, etc) and, within each group, each unit is ranked according to its membership in an adjacent cluster. Units in *group 1* are ranked in ascending order of their memberships in *cluster 2*, units in *group 2* are ranked in ascending order of their memberships in *cluster 3*, units in *group 3* are ranked in ascending order of their memberships in *cluster 4*, and units in *group 4* are ranked in descending order of their memberships in *cluster 3*. Triangles, *cluster 1;* stars, *cluster 2;* squares, *cluster 3;* circles, *cluster 4*. *Bottom*: disparity tuning curves; each *point* along the *abscissa* corresponds to the unit whose membership value is plotted above, and its disparity tuning curve is shown in a color-coded form in accordance with the scale shown at the right (increases in discharge rate, in imp/s, are shown in red and decreases are shown in blue) with the amplitude of the disparity steps represented along the *ordinate*.

vergence data for both animals: compare the ○ (summed activity) and * (vergence) plotted in Fig. 6*A* (*monkey N*, $r^2 = 0.98$) and Fig. 6*B* (*monkey Q*, $r^2 = 0.93$). Inverting the sign of the curves with negative slopes around zero disparity was critical for achieving such good fits. Thus failure to invert those curves (before fitting) decreased the r^2 to 0.12 for *monkey N* and to 0.21 for *monkey Q*, and the best fits now showed little resemblance to the tuning curves for vergence: see Fig. 6, *C* and *D*. None of the single units had a tuning curve that matched the vergence as well as these population responses did: even with sign inversion, the highest r^2 for any individual unit was 0.92 for *monkey N* (mean r^2: 0.49) and 0.86 for *monkey Q* (mean r^2: 0.48). It is of interest that the disparity tuning curves for the summed neuronal activity, like those for the vergence eye movements, were well fit by Gabor functions, r^2 being 0.95 and 0.94 for *monkeys N* and

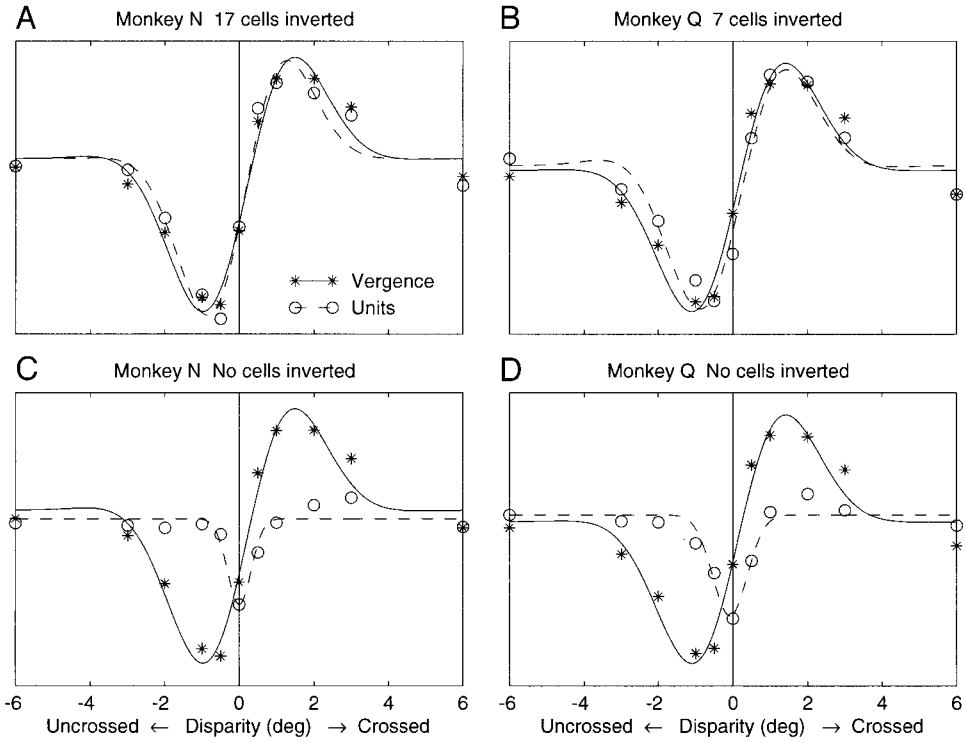

FIG. 6. Spatial coding of vergence by populations of cells (correlated stimuli). Disparity tuning data for vergence (*) and for the least-squares, best-fit summed activity (○) for *monkey N* (*A* and *C*) and *monkey Q* (*B* and *D*). For the *top plots* (*A* and *B*), the individual unit curves with negative slopes around zero disparity were inverted before summing and for the *bottom plots* (*C* and *D*), no curves were inverted. Curves are the least-square, best-fit Gabor functions for the vergence data (—) and the summed activity data (- - -).

Q, respectively: see - - - plotted in Fig. 6. The parameters for these fits are included in Table 1.

These data indicate that the disparity tuning curve for the vergence eye movements was effectively encoded in the aggregate activity of all the disparity-sensitive cells recorded in MST. The question arises as to the relative contributions of the four groups identified by the fuzzy cluster analysis. We addressed this issue by examining the effect of excluding each of the groups of units on the goodness of the fit between the disparity tuning curves for the summed activity and for the vergence eye movements when, as usual, gain and *y* offset were free parameters. Removing all of the cells in any one of the four groups always increased the MSE (and thus reduced r^2) for *monkey N*. We examined the statistical significance of this finding using a bootstrap algorithm (Efron and Tibshirani 1991), which involved computing the MSE after randomly excluding equivalent numbers of units, using 10,000 random combinations of cells. This indicated that the probability that excluding a random sample of cells would have an impact equal to that when any one of the four groups of cells was removed was always <0.008. This implies that units in all four of the groups made a significant positive contribution to the population response of *monkey N*. However, the data for *monkey Q* were less consistent and only the removal of the tuned far *group 1* resulted in a significant worsening of the best fit ($P < 0.001$, bootstrap algorithm).

Discrete coding? To further examine the possibility of the discrete coding of vergence by a subpopulation of MST cells, we used a genetic algorithm to determine if there was a subset of cells whose activity when summed together resulted in a disparity tuning curve that fitted the disparity tuning curve for vergence as well as, or better than, that from summing the entire population. We again treated the data from each animal separately and inverted the sign of those disparity tuning curves with negative slopes in the disparity range $\pm 1°$. We ran the algorithm 100 times on each data set, and the number of cells that survived to the last generation each time varied from 19 to 28 (of 49) for *monkey N* and from 7 to 14 (of 31) for *monkey Q*. The MSEs for the summed activity of the last generations of cells were smaller than those when the whole population was summed, by a factor of 58–107 for *monkey N* ($r^2 = 0.9995$ for the subset giving the very best fit) and by a factor of 16–24 for *monkey Q* (highest $r^2 = 0.9973$). Clearly, all of the subsets of cells selected by the genetic algorithm gave much better fits to the vergence data than did the whole populations, and the very best fits (i.e., the best of each 100) were almost indistinguishable from the targeted vergence data: see Fig. 7, *A* and *B*. Every one of the four groups contributed to these very best fits, though *group 2* always contributed fewest: for *monkey N*, each group contributed 7, 2, 5, and 5 cells (in order from *group 1* through to *group 4*), and for *monkey Q*, these numbers were 4, 1, 2, 2. The special importance of *groups 1* and *4* was evident when the genetic algorithm was run after excluding all of the curves selected from either of those groups: the very lowest MSE (i.e., the lowest when the algorithm was run 100 times) increased by a factor of 124 and 26, respectively, for the data from *monkey N*, and by a factor of 9 and 2, respectively, for the data from *monkey Q*. In contrast, excluding the cells selected from *group 2* had no significant impact on the very lowest MSE in either monkey, and excluding those selected from *group 3* had a significant impact only for *monkey N* (increasing the very lowest MSE by a factor of 12).

Temporal coding? In view of the fact that the disparity tuning curves for the summed activity of the whole population of cells showed a good correlation with the disparity tuning curves for vergence (spatial coding), we also attempted to determine if the summed activity elicited by a given disparity stimulus conveyed information about the time course of the vergence responses (temporal coding). As before, we first

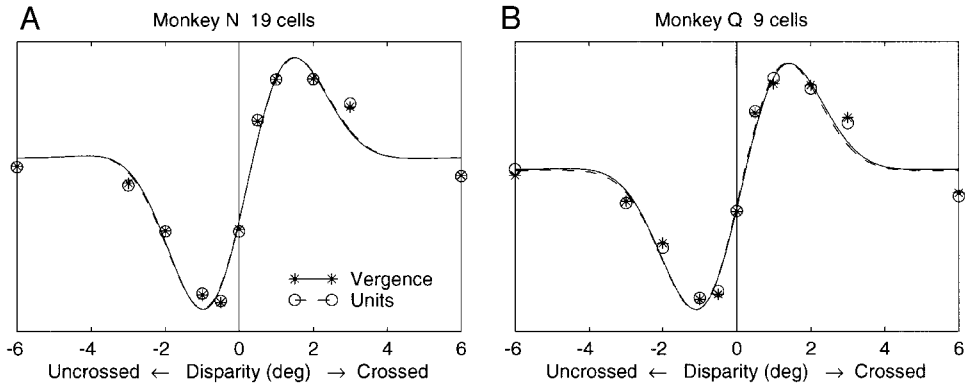

FIG. 7. Spatial coding of vergence by the subsets of cells selected by the genetic algorithm (correlated stimuli). Disparity tuning data for vergence (*) and for the summed activity of the subsets of cells (selected by the genetic algorithm) that gave the very best fits to the vergence data (○) for *monkey N* (*A*) and *monkey Q* (*B*). Curves are the least-squares best-fit Gabor functions for the vergence data (—) and for the summed unit data (- - -).

inverted the contributions of those cells whose disparity tuning curves had negative slopes in the disparity range ±1° and treated the data from *monkeys N* and *Q* separately. It was soon apparent that noise was a major limiting factor here, but the summed discharge profiles elicited by some disparity stimuli strongly resembled the associated vergence velocity profile, and we obtained least-squares best fits, allowing gain, *x* offset (time delay), and *y* offset to be free parameters. Samples of two such fits can be seen in Fig. 8, *A* and *B*, which shows the time course of the average vergence velocity response elicited by 1° crossed-disparity steps for *monkeys N* and *Q* together with the best-fit discharge-rate profiles using the summed activity of all cells from each animal. The fit was clearly better for *monkey N* ($r^2 = 0.98$; $n = 49$) than for *monkey Q* ($r^2 = 0.87$; $n = 31$), possibly in part because of the difference in the numbers of cells. The *x* offset that yielded these best fits was −18 ms for *monkey N* and −23 ms for *monkey Q*. In both cases, the peak in the relationship between r^2 and the time delay was sharply convex (Fig. 8, *C* and *D*), indicating that these time delays provide a reliable estimate of the time interval by which the neuronal population response preceded the vergence response.

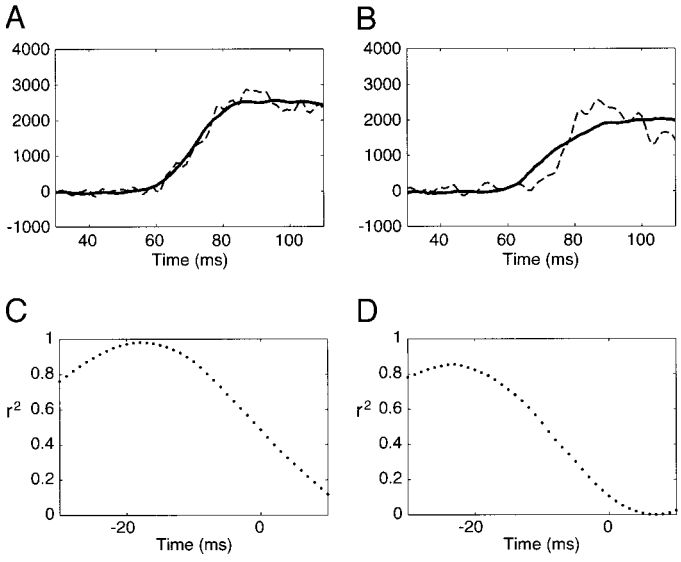

FIG. 8. Temporal coding of vergence velocity by populations of cells (correlated stimuli). *A* and *B*: plots of average vergence velocity (—) and the least-squares, best-fit summed activity (- - -) over time for *monkey N* (*A*) and *monkey Q* (*B*) in response to 1° crossed-disparity steps; gain and offset were free parameters for these fits, which were obtained for a range of *x* shifts (time shifts, at 1-ms intervals), and the data shown are for the *x* shifts that gave the very best fits. *C* and *D*: the dependence of the goodness of these best fits (r^2) on the time shift is plotted for *monkey N* (*C*) and for *monkey Q* (*D*); a negative time shift indicates that the summed activity has been moved forward in time with respect to the vergence response; both peak at negative values (−18 ms in *A*, −23 ms in *B*), indicating that the summed activity preceded the vergence response and had to be moved forward to get the very best fit.

However, r^2 for the least-squares best fits exceeded 0.9 for only 8/20 data sets, most of these (7) involving responses to crossed disparity steps, and the *x* shifts that gave these fits ranged from −15 to −22 ms (mean ± SD, 17.6 ± 2.3 ms): see Table 2 (values listed under "all cells"). The fits were even worse when we summed only the activity of the cells that had been selected with the genetic algorithm: see Table 2 (values listed under "GA"). Visual inspection revealed that the summed discharge profiles giving poor fits were generally noisy, often before the onset of the disparity-driven response, indicating the existence of appreciable noise unrelated to the disparity stimulus. To our surprise, the summed discharge profiles that gave good fits always included individual profiles that varied widely. For example, the summed discharge profile that included the data in Fig. 2 accounted for >93% of the stimulus-induced variation in the associated vergence velocity profile ($r^2 = 0.933$), despite the obvious variability among the individual cells in the latency and time course of the (averaged) discharge profiles. Thus we suggest that the paucity of good fits was due in large part to the inadequacy of our data samples: on the one hand, the relatively small number of responses averaged for each stimulus condition and, on the other, the relatively small numbers of cells recorded from any given monkey. Another important factor here was that not all units were active for all stimuli, further limiting the number of discharges contributing to a given temporal profile. In summary, although noise problems restricted the amount of useful data, the temporal profile of the summed activity associated with some disparity steps showed a reasonably good fit to the vergence velocity profile.

Initial vergence responses to disparity steps (anticorrelated patterns)

The vergence responses elicited by disparity steps applied to anticorrelated patterns were recorded from two animals (*monkeys N* and *Q*). As previously reported by Masson et al. (1997), these vergence responses were comparable in latency with those produced by the same steps applied to correlated patterns but were often in the opposite direction. This is apparent from the sample data from *monkey N* shown in Fig. 9, which has a layout identical to Fig. 1. Thus the disparity tuning curves

TABLE 2. *Temporal coding of vergence velocity by populations of cells: coefficients of determination (r^2)*

	Correlated Stimuli				Anticorrelated Stimuli	
	Monkey N		*Monkey Q*		*Monkey N*	*Monkey Q*
Disparity Step, °	All cells	GA	All cells	GA	All cells	All cells
−6	0.917 (−17)	0.854 (−19)	0.893 (−18)	0.410 (−23)	0.634 (−21)	0.266 (23)
−3	0.894 (−18)	0.776 (−18)	0.536 (−14)	0.417 (−15)	0.684 (−21)	0.754 (−26)
−2	0.486 (−10)	0.143 (−30)	0.597 (−30)	0.530 (−15)	0.862 (−14)	0.942 (−27)
−1	0.785 (−17)	0.765 (−27)	0.630 (25)	0.626 (−21)	0.608 (−15)	0.939 (−25)
−0.5	0.938 (−17)	0.859 (−23)	0.584 (−28)	0.695 (−27)	0.664 (−30)	0.817 (−20)
0.5	0.981 (−15)	0.949 (−18)	0.892 (−21)	0.866 (−17)	0.841 (−23)	0.612 (−19)
1	0.980 (−18)	0.940 (−20)	0.874 (−23)	0.894 (−17)	0.837 (−21)	0.744 (−15)
2	0.933 (−18)	0.888 (−21)	0.942 (−22)	0.845 (−15)	0.791 (−19)	0.518 (−16)
3	0.947 (−15)	0.930 (−21)	0.912 (−19)	0.755 (−4)	0.672 (−18)	0.456 (21)
6	0.789 (−16)	0.790 (−19)	0.308 (−7)	0.110 (−14)	0.723 (−18)	0.514 (−18)

The temporal profiles of the summed activity were fitted to the temporal profiles of the associated mean vergence velocity for *monkeys N* and *Q*, and the coefficients of determination were computed. Data are based on either the whole population of curves (all cells) or the subsets previously selected by the genetic algorithm (GA) using the disparity-tuning data. Gain and *y* offset were free parameters for these least-square, best fits, which were obtained for a range of *x* shifts, and the data listed are for the *x* shifts that gave the best fits. These optimal time shifts (in ms) are shown in parentheses. Two examples of these best fits are shown in Fig. 8, *A* and *B*.

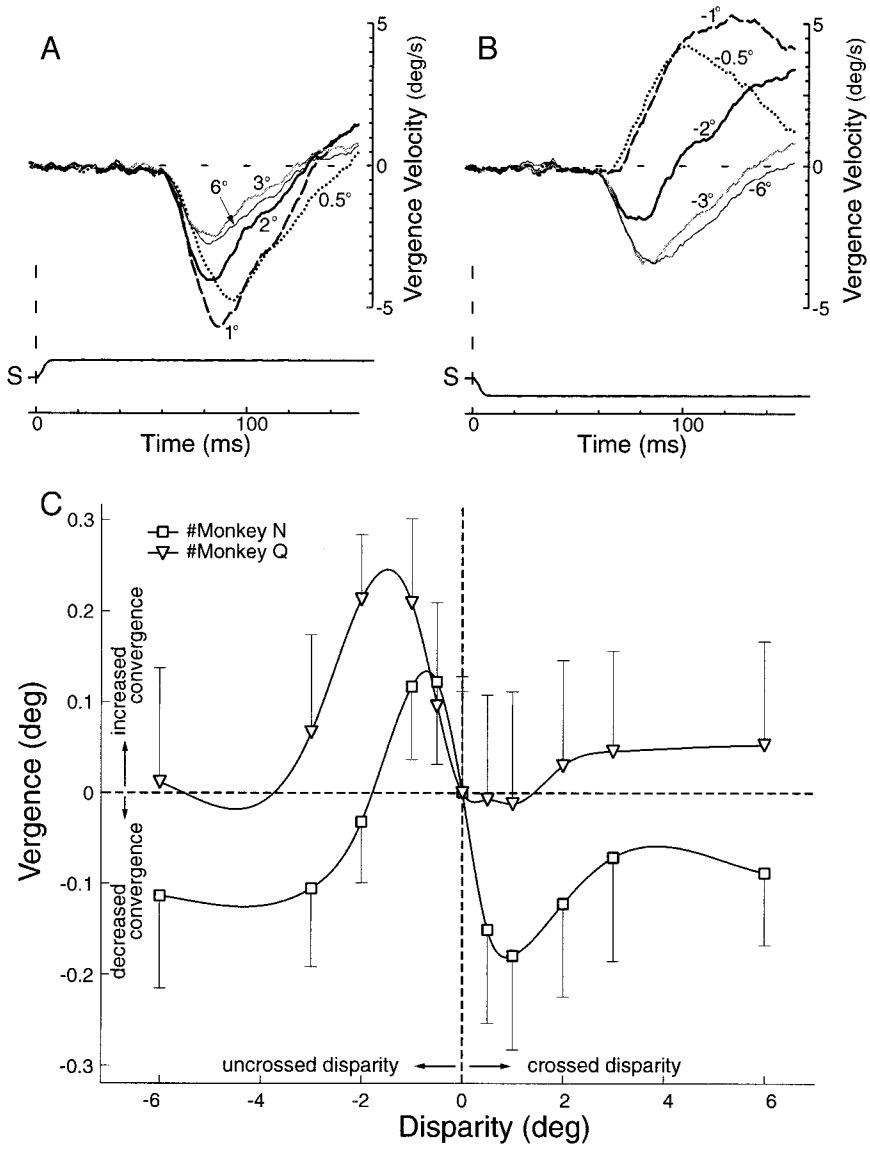

FIG. 9. Vergence responses: dependence on the amplitude and direction of the disparity stimulus (anticorrelated stimuli). Traces in *A* show vergence velocity over time in response to crossed disparity steps, and traces in *B* show the same for uncrossed steps. *C*: the mean change in vergence position in degrees, measured over the 60-ms period starting 50 ms after the disparity step), is plotted against the magnitude of the step in degrees, for each of the 2 monkeys: disparity tuning curves. All conventions as for Fig. 1.

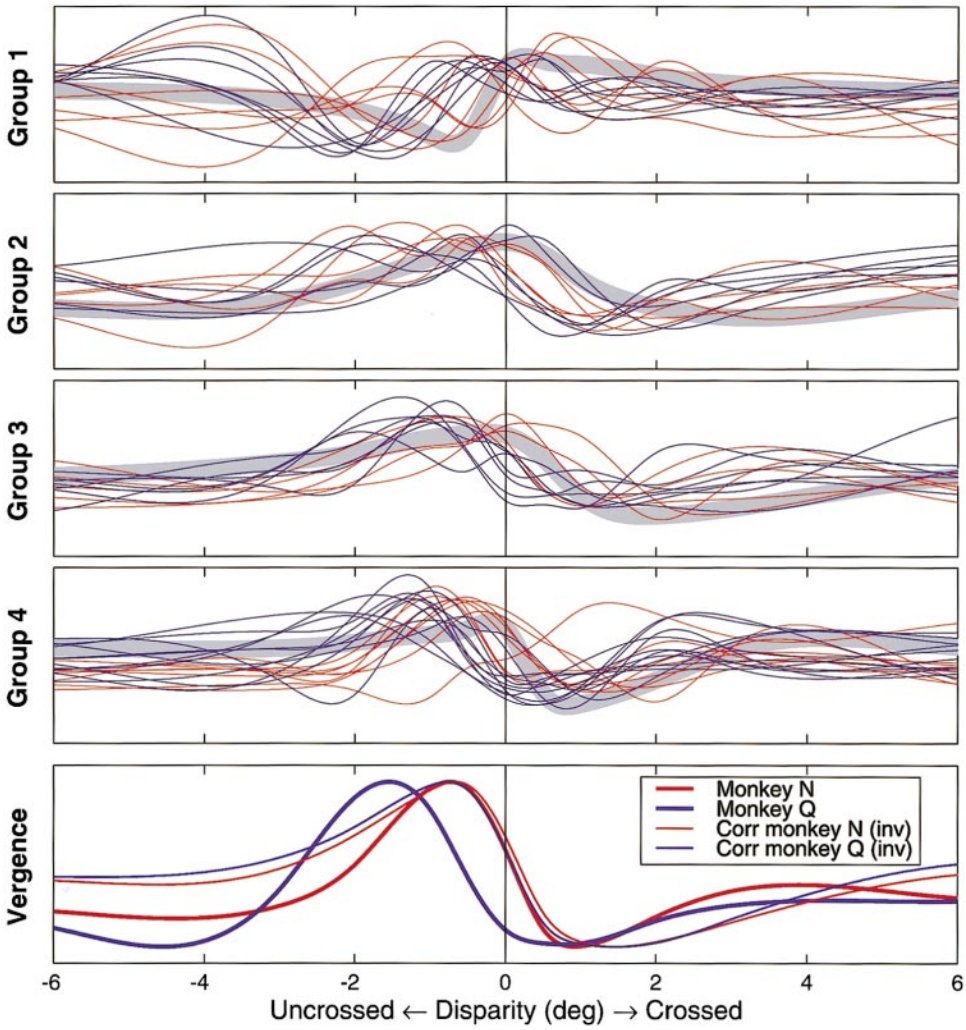

FIG. 10. Disparity tuning curves for the individual cells (anticorrelated stimuli). *Top 4 graphs*: mean change in discharge rate over the 60-ms period starting 40 ms after the disparity step plotted against the magnitude of the disparity step for the data from 2 monkeys (*N*, data in red; *Q*, data in blue); curves are normalized and each has been assigned to 1 of 4 groups (the curve for a given cell being assigned to the group to which that cell's *curve* was assigned in Fig. 4); *thick gray traces* in each of the 4 graphs show the (inverted) median tuning curves obtained for the same groups of cells with correlated stimuli. *Bottom graph*: *thick colored traces*, the disparity tuning curves for the associated mean vergence responses of the 2 monkeys; *thin colored traces*, the (inverted) vergence tuning *curves* obtained from the same 2 monkeys with correlated stimuli. Traces are spline interpolations (see legend of Fig. 1 for details).

showed a negative slope in the immediate vicinity of zero disparity, though it is clear that the curve for *monkey Q* is shifted appreciably to the left of the curve for *monkey N*. The peak-to-peak amplitudes of the (interpolated) disparity tuning curves for the anticorrelated data were smaller than those for the correlated data: by 33% for *monkey N* and by 44% for *monkey Q*. The tuning curve data were well fit by a Gabor function (r^2 was 0.99 and 0.98 for *monkeys N* and *Q*, respectively), and the parameters of the least-squares best fits are listed in Table 1: see also the continuous lines plotted in Fig. 12, *A* and *B*. Of particular interest among the parameters of the Gabor functions is the difference in the phase of the cosine terms for the correlated and anticorrelated data: 177° for *monkey N* and 169° for *monkey Q*. This reinforces the impression that, to a first approximation, the disparity tuning curves obtained with anticorrelated stimuli were inverted versions of those obtained with correlated stimuli. The above-mentioned shifts, however, are evident in the *x* shifts of the best-fit Gabor functions, which differed for the correlated and anticorrelated data, especially for *monkey Q*: this parameter was always zero for the correlated data and negative for the anticorrelated data (Table 1). In fact, the tuning curves obtained with anticorrelated stimuli were quite well fitted by the tuning curves obtained with correlated stimuli when the latter were inverted provided that the *x* shift was a free parameter (in addition to gain and *y* offset): r^2 values for the least-squares best fits for *monkeys N* and *Q* were 0.86 and 0.97, respectively, and the corresponding *x* shifts were 0.05 and 0.61°.

Initial neuronal responses to disparity steps (anticorrelated patterns)

Neurons that were still well isolated after we had finished recording their responses to disparity steps applied to correlated patterns were then recorded while the same steps were applied to anticorrelated patterns. The activity of 56 MST neurons in three hemispheres of two monkeys was so recorded (25/49 units from *monkey N* and 31/31 units from *monkey Q*), and all gave significant responses to the anticorrelated stimuli ($P < 0.005$, 1-way ANOVA). Neuronal response latencies to anticorrelated stimuli were roughly comparable to those obtained with correlated stimuli.

DISPARITY TUNING CURVES OF SINGLE CELLS. Figure 10 shows the normalized disparity tuning curves for all 56 units whose responses to anticorrelated stimuli were recorded, and is organized like Fig. 4, with cells placed in the same four groups. That is, cells that were in *group 1* in Fig. 4 were also placed in *group 1* in Fig. 10, and so forth. Again, we fitted the disparity tuning curves of the individual cells to the tuning curves for

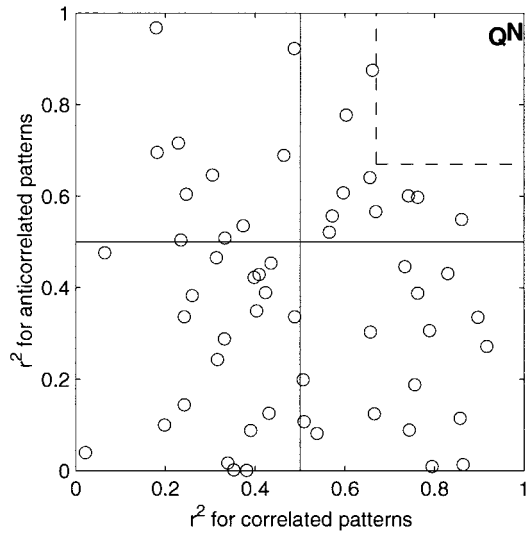

FIG. 11. Coding of vergence by individual cells (correlated and anticorrelated stimuli). The disparity tuning curves of the individual cells were fitted to the disparity tuning curves of the associated vergence responses and r^2 values for the least-squares, best fits were computed. This graph plots the r^2 values for the data obtained with correlated patterns against those obtained with anticorrelated patterns. No cell had r^2 values that exceeded 0.67 for both stimuli (indicated by the - - -). Also shown (indicated by their identifying letters), are the r^2 values for the fits between the summed activity and the vergence responses for the 2 monkeys, *N* and *Q* (from which all the unit data plotted here were obtained).

vergence (gain and *y* offset, free parameters) and assumed that the contribution of any given cell to the vergence response would always have the same sign regardless of the stimulus used to drive it. Accordingly, cells whose contributions had been inverted for the earlier analysis of the correlated data were again inverted here. (For *monkey N*, 7/25 curves were inverted and for *monkey Q*, 7/31 curves were inverted, all in *group 1*.) As for the correlated data, the least-squares best fits for the anticorrelated data ranged widely (r^2: mean, 0.39; range, 0–0.97) and some were clearly good enough to be considered vergence-related rather than disparity-related. However, no single cell had responses that fitted the vergence responses obtained with both correlated and anticorrelated stimuli with an r^2 value >0.67: Fig. 11 shows a plot of the individual r^2 values obtained with correlated stimuli against those obtained with anticorrelated stimuli, and there is a relative paucity of cells in the upper right quadrant, which is where pure vergence-encoding cells would be expected.

Comparing Fig. 10 with Fig. 4 indicates that, within each group, the anticorrelated data have much more scatter than the correlated data: cells that had similar tuning curves with correlated stimuli (and so were assigned to the same group) often had very different tuning curves with anticorrelated stimuli. Some of this scatter for disparities between 3 and 6° (where there are no data points) comes from the spline interpolation: see particularly the *group 1* curves with uncrossed disparities. Further, some of the scatter in Fig. 10 might have resulted from differences in the responses of the two monkeys: for example, in *group 4*, the peaks for *monkey Q* (shown in blue) are often to the left of those for *monkey N* (shown in red), and a similar (weaker) trend is evident in *group 3*. As we saw earlier, the tuning curve for the vergence eye movements of *monkey Q* is shifted to the left of that for *monkey N* (see also the thick colored traces in the *bottom plot* of Fig. 10), and we sought to investigate these *x* offsets further.

We saw in the preceding text that, to a first approximation, the vergence responses to correlated and anticorrelated stimuli had opposite signs, so that the disparity tuning curve for the vergence data obtained with one stimulus resembled the inverse of the other, except for an *x* offset that was greater in *monkey Q* than *monkey N*. We attempted to determine if this also applied to the single-unit responses, i.e., whether, for each cell, the response to anticorrelated patterns closely resembled the inverse of the response to correlated patterns. There is some hint of this in Fig. 10, which includes, in addition to the individual disparity tuning curves for the responses to anticorrelated patterns, plots of the inverse of the group median responses to correlated patterns: see the thick gray lines in each group.[3] Thus the curves in *groups 2–4*, which show more consistency than *group 1*, often have a general form that resembles the inverted median but shifted horizontally. When the individual tuning curves obtained with anticorrelated stimuli were fitted to the inverted median responses obtained with correlated stimuli (with gain, *y* offset and *x* offset as free parameters), many of the fits were quite good: for *monkeys N* and *Q*, mean r^2 values were 0.61 and 0.71 and mean *x* shifts were 0.31 and 0.73°, respectively. This difference in the *x* shifts of the data from the two monkeys (0.42°) was statistically significant ($P = 0.02$, 1-way ANOVA) and corresponded roughly with the difference in the *x* shifts reported in the preceding text when the inverted tuning curve for the vergence responses to correlated stimuli was fitted to the tuning curve for the vergence responses to anticorrelated stimuli (0.56°). The need for this latter shift is clear from the *bottom plot* in Fig. 10, which includes the inverted disparity tuning curves for the vergence responses to correlated stimuli (thin colored traces). Note that we did not attempt to fit the response of each cell to anticorrelated stimuli with the inverse of the response of the same cell to correlated stimuli because of the *x* offsets, which would require the pairing of actual data points from one tuning curve with interpolated data points from the other. As the use of interpolated data points makes the result sensitive to the quality of the fit, we used the median because, being computed over a number of cells, it smoothes out the oscillations (compare Fig. 10 with Fig. 4).

DISPARITY TUNING CURVES FOR (SUMMED) POPULATIONS OF CELLS. We saw earlier that, when correlated patterns were used, the disparity tuning curves for the summed activity of the population of cells matched the tuning curves for the associated vergence responses quite well. We now sought to determine if the same was true for the data obtained with anticorrelated stimuli. We again summed all of the raw—that is, nonnormalized—tuning curves for each of the two monkeys separately and then determined how well these population responses fitted their respective tuning curves for vergence when gain and *y* offset were the only free parameters. We assumed that the contribution of any given cell to the vergence response would always have the same sign regardless of the stimulus used, and cells whose contributions had been inverted for the earlier analysis of the correlated data were again inverted here. Figure

[3] The medians provided a fairly good representation of the correlated data: mean r^2 values between individual tuning curves and the median for groups 1–4 were 0.74, 0.79, 0.82, and 0.78, and r^2 exceeded 0.4 for 100/102 cells.

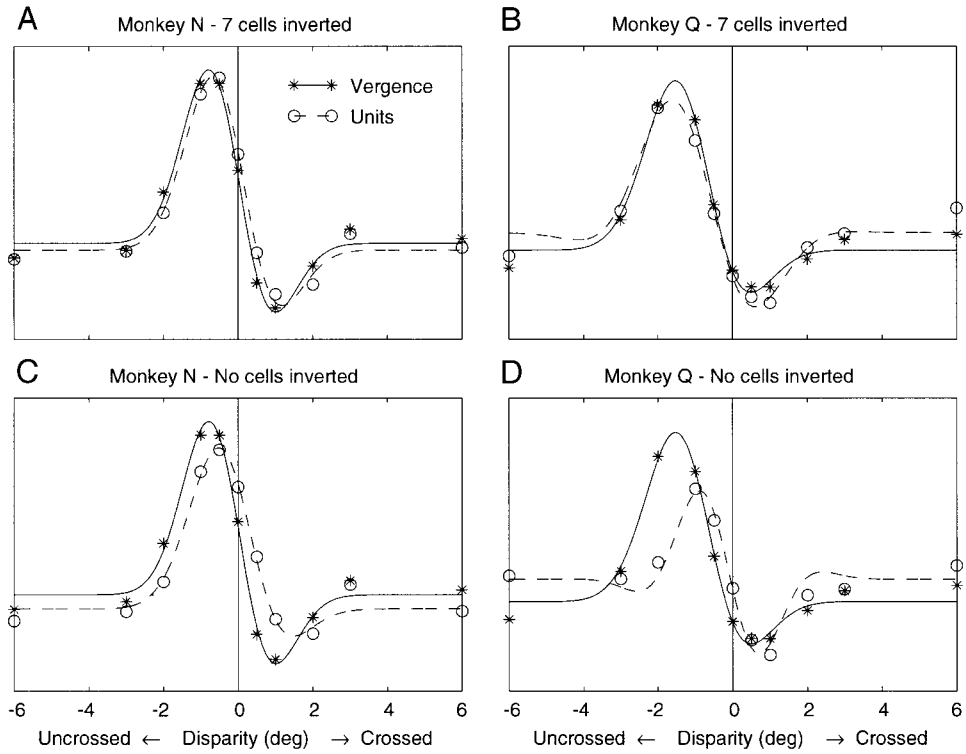

FIG. 12. Spatial coding of vergence by populations of cells (anticorrelated stimuli). Disparity tuning data for vergence (*) and for the least-squares, best-fit summed activity (○) for monkey N (*A*, *C*) and monkey Q (*B*, *D*). For the upper plots (*A*, *B*), the curves with negative slopes around zero disparity with correlated stimuli were inverted before summing, and for the lower plots (*C*, *D*) no curves were inverted. *Curves* are the least-squares best-fit Gabor functions for the vergence data (——) and the summed activity data (– – –).

12, *A* and *B*, shows that the disparity tuning data for the summed activity (○) again matched those for the associated vergence responses (*) quite well (r^2: 0.96 and 0.95 for *monkeys N* and *Q*, respectively). Once again, failure to invert the contributions of the relevant cells before fitting led to significantly worse fits (r^2: 0.77 and 0.57 for *monkeys N* and *Q*, respectively), though the effects on the fits were not as dramatic as reported above for the data obtained with correlated patterns. In the case of *monkey N*, this is perhaps in part because a somewhat smaller proportion of the curves obtained with anticorrelated stimuli were inverted: 28% (7/25), compared with 35% (17/49) of those obtained with correlated stimuli. There was one cell from *monkey N* whose tuning curve matched the vergence as well as the curve for the summed activity did ($r^2 = 0.97$) but the mean r^2 for all cells from this monkey was only 0.47. None of the single units from *monkey Q* had tuning curves matching its vergence responses as well as its summed activity did and the highest r^2 for any individual unit from this monkey was 0.88 (mean r^2: 0.32).

The curves in Fig. 12 are the best-fit Gabor functions and (again) clearly provide a good representation of the disparity tuning curves for both vergence and the summed activity, with r^2 values of 0.99 and 0.99, respectively, for *monkey N*, and 0.98 and 0.96, respectively, for *monkey Q*. The parameters of these Gabor functions are listed in Table 1. It was pointed out earlier that the best-fit Gabor functions for the vergence eye-movement data obtained with correlated and anticorrelated stimuli differed in the phase of their cosine terms by close to 180°, and it is apparent from Table 1 that the same was true for the summed activity data: the phase difference was 178° for *monkey N* and 170° for *monkey Q*. We further noted earlier that the best-fit Gabor functions for the correlated and anticorrelated vergence data differed also in their *x* offset, and Table 1 indicates that the *x* offset for the summed activity followed a very similar pattern: the *x* offset was always zero for the correlated data, and negative for the anticorrelated data. Additionally, these negative values for the anticorrelated data differed markedly for the two monkeys and by an amount that was comparable for the vergence and summed-activity data. Interestingly, similar trends were evident in the *y*-offset (Table 1). In sum, the best-fit Gabor functions provide a good representation of the disparity tuning curves for both the summed activity and the associated vergence response, and even document their shared idiosyncrasies.

It is interesting that the differences in the anticorrelated data of the two monkeys were significant enough that the summed activity from *monkey N* gave a very poor fit to the vergence data from *monkey Q* ($r^2 = 0.29$), and vice versa ($r^2 = 0.35$). Of course, with correlated stimuli, the vergence behavior of the two monkeys was very similar, and it is therefore not surprising that, with these stimuli, the vergence data from one monkey were well fit by the summed activity of the other: the summed activity from *monkey N* fit the vergence data from *monkey Q* with an r^2 value of 0.96, and the summed activity from *monkey Q* fit the vergence data from *monkey N* with an r^2 value of 0.93.

When the genetic algorithm was applied to the disparity tuning curves of the cells examined with the anticorrelated stimuli (inverting the signs of some curves as usual), the number of cells that survived to the last generation varied from 13 to 21 (of 31) for *monkey Q* and was always 13 (of 25) for *monkey N* (though not always the same 13 cells). Once again, the summed activity of the subsets of cells selected by the genetic algorithm always gave extremely good fits to the vergence data: when the algorithm was run 100 times, the MSEs for the last generations of cells were smaller than those for the whole population by a factor of 28–35 for *monkey N* (highest $r^2 = 0.9985$) and 22–45 for *monkey Q* (highest $r^2 = 0.9990$). Further, as for the correlated data, every one of the four cell

groups contributed to the very best of the 100 fits, and *groups 1* and *4* always contributed most: for *monkey N*, each group (in order, from 1 to 4) contributed 5, 2, 3, and 3 cells, and for *monkey Q* these numbers were 5, 2, 5, 7.

Each of the temporal profiles of the summed activity elicited by a given disparity step was fitted to the temporal profile of the associated vergence-velocity response, treating the data from each animal separately, inverting some curves as usual, and allowing gain, *x* offset (time delay), and *y* offset to be free parameters. Unfortunately, visual inspection revealed that many of these temporal profiles had low signal-to-noise and r^2 exceeded 0.9 for only 2/20 fits (Table 2).

DISCUSSION

About 20% of the cells that we recorded in MST responded to disparity steps, their discharges often commencing at short latency and preceding the associated vergence eye movements. The depth tuning of many cells resembled that previously reported for cells in various regions of the visual cortex: only a few cells in our sample shared the depth tuning of the associated vergence responses and then only for the data obtained with one of the two kinds of disparity stimuli—correlated or anticorrelated patterns—and never for both. Thus we were surprised to find that, for both stimuli, merely summing the disparity tuning curves of all the cells sampled in a given animal resulted in a curve that closely approximated the disparity tuning curve for that animal's vergence responses. Further, for some disparity stimuli, the time course of the initial changes in the summed activity closely mirrored the time course of the associated changes in vergence velocity: there was little hint of this in the discharge profiles of the individual cells. These data suggest that the discharges of the individual cells each encode some limited aspect(s) of the disparity stimulus and/or vergence motor response, whereas the summed activity of the population (faithfully?) encodes the entire vergence velocity response. We will discuss the information carried by single cells and by the population separately.

Information coding at the single-cell level

DISCRETE GROUPS OR A CONTINUUM? Most of the previous studies of the disparity-selective neurons in the monkey cortex used the classification schemes of Poggio and Fischer (1977) or Poggio et al. (1988), whose recordings were all made in the visual cortex (V1, V2, and V3). The later study divided disparity tuning curves into two major groups, tuned and reciprocal, each with subgroups based on their specific preferences for images located in front of (tuned near, near), within (tuned zero, tuned inhibitory), or behind (tuned far, far) the plane of fixation. These groupings have proven very useful for describing the disparity-selectivity of cortical neurons especially in the various regions of the visual cortex for which they were originally devised, although their applicability to higher levels of cortex is less clear. For example, in the study of Maunsell and Van Essen (1983a) on MT, which projects directly to MST (Maunsell and Van Essen 1983b; Ungerleider and Desimone 1986), most of the disparity-sensitive cells fell into one of the four categories, near, far, tuned zero, and tuned inhibitory, but almost 12% "were on a border between two classes." Roy et al. (1992) reported that most of the disparity-sensitive cells in the dorsomedial region of MST (referred to as, MSTd) were in the near or far categories, a few were tuned, and 17% had hybrid curves with features of both tuned and reciprocal curves. Eifuku and Wurtz (1999) reported that disparity-sensitive cells in the lateral region of MST (referred to as, MSTl) were broadly tuned and did not fit the classification scheme of Poggio et al. These workers also suggested that there was "a continuum of disparity tuning functions" in MSTl, some curves having clear peaks and others not. Poggio et al. (1988) had earlier suggested that four of the classes in visual cortex (near, tuned near, tuned far, far) might form a continuum, and DeAngelis and Newsome (1999) reported that the disparity preferences of cells in MT ranged from near to far without any major discontinuities, mapping smoothly across the surface of the cortex. Further, Gnadt and Mays (1995), in a study of the depth dependence of saccade-related neurons in the lateral bank of the intraparietal sulcus (LIP), which has reciprocal connections with MST (Andersen et al. 1990; Blatt et al. 1990; Ungerleider and Desimone 1986), also found that not all of the tuning curves matched one of the groups of Poggio et al. and concluded that there was "little basis for categorical parcellation into separate functional classes."

In our study, disparity tuning curves were sorted into categories without regard for any previous classification schemes, using an objective method based on the fuzzy c-means clustering algorithm (Bezdek 1981). Nonetheless, this approach identified four potential groups of disparity-selective cells that resembled four of the (6) classes recognized by Poggio et al. (1988): tuned far, tuned inhibitory, near, tuned near. In common with previous studies of MT and MST, however, we found that the tuning for disparity was often quite broad compared with that in visual cortex. Thus although the cells in our sample with the narrowest peaks—those in *groups 1* and *4*—were likened to the cells with the narrowest (nonzero) peaks in the grouping of Poggio et al.—the tuned far and tuned near categories—the fact is that the narrowest peaks in our study were often so broad and skewed that these cells would probably have been placed in the far and near categories had they been recorded in striate cortex. It is also important to remember that we decided on four clusters only after finding that this was the largest number that gave consistent groupings when the algorithm was repeated. This should not be taken to imply that there are four discrete groups of cells: The membership values indicated that, with the possible exception of *group 1*, our groups were fuzzy and that *groups 2–4* might well form a continuum. Indeed, using the membership values in the nondefining clusters to rank order the tuning curves within each group indicated that there was little evidence of discontinuities at the boundaries between the three groups (Fig. 4). In addition, the hierarchical fuzzy algorithm, which attempts to find the maximum number of discrete clusters that can be justified by the data and continues to merge clusters until a firm criterion for separation between clusters is met, merged our *groups 2–4* and hence recognized only two clusters as discrete. However, it is still possible that, with more data, the hierarchical fuzzy algorithm would recognize more clusters, and so the question of whether *groups 2–4* are discrete entities or parts of a continuum must remain open for the present. Regardless, we do feel that it is useful to recognize four groups for descriptive purposes. Thus the tuned inhibitory cells of *group 2*, for example, are clearly very different from the tuned near cells of *group 4*.

INITIAL OPEN-LOOP RESPONSES. A somewhat unusual aspect of our study is that it dealt exclusively with the so-called open-loop neural responses that occur at short latency (40–100 ms after stimulus onset), before closure of the disparity-feedback loop. In a recent study of neurons in the frontal eye fields (FEF), which receive direct projections from MST (Schall et al. 1995), disparity stimuli were presented suddenly against a previously blank background, and disparity tuning curves based on the responses during an early period (50–200 ms after stimulus onset) could have a very different shape from those based on the responses during a later period (300–1,000 ms after stimulus onset) (Ferraina et al. 2000).

SENSORY AND/OR MOTOR? In previous studies, the disparity-selective activity in MT and MST was assumed to be purely visual in origin, and, insofar as it was linked to vergence eye movements at all, it was solely as a possible source of visual feedback guiding binocular alignment (Maunsell and Van Essen 1983a; Roy et al. 1992). In our experiments, unfortunately, the sensory and motor events were so time-locked that we were unable to determine whether the responses were more closely synchronized to one or to the other. Thus in our experiments, the discharges of individual neurons might encode some aspect of the stimulus (disparity) and/or the associated motor response (vergence).

The complete lack of units in the upper right of Fig. 11 indicates that none of the individual cells provided a complete representation of vergence to rival that seen at the level of the summed population and argues against there being a homogeneous subset of cells that each encode the entire vergence response: at best, the discharges of any given cell can convey only some limited aspect of the vergence response. This view is reinforced by the finding that the subsets of cells selected by the genetic algorithm were always drawn from all four groups: a good representation of the vergence response required a heterogeneous collection of cells, each cell contributing to a different aspect of the response. The finding that cells that had similar tuning curves with correlated stimuli (and so were assigned to the same group) often had very different tuning curves with anticorrelated stimuli (compare Figs. 4 and 10) indicates that the vergence eye movements alone did not fully account for the discharges. This is also evident from the scatter in Fig. 11, which indicates that, across the population of cells, the goodness of the fit between neuronal and vergence responses for the correlated and anticorrelated data were unrelated ($r^2 = 0.006$; slope $= -0.08$). Thus the type of disparity stimulus used to elicit the vergence responses also influenced the shape of the individual tuning curves independent of any effect on the vergence itself. It is entirely possible that some (or all) cells carry mixed sensory and motor signals, an arrangement for which there is a precedent in MST: some neurons that discharge in relation to optic flow and pursuit tracking show both retinal and extraretinal influences, the latter encoding some aspect of the tracking eye movements (Bradley et al. 1996; Newsome et al. 1988; Page and Duffy 1999; Sakata et al. 1983; Shenoy et al. 1999; Squatrito and Maioli 1996, 1997; Thier and Erickson 1992). These data notwithstanding, none of our data rule out the possibility that, as generally supposed, the disparity sensitivity of individual MST cells is purely sensory and hence independent of the associated motor (vergence) responses per se.

Information coding at the population level

The summed activity of the disparity-sensitive cells in MST carries considerable information about the initial vergence eye movements associated with a sudden change in disparity. We examined this population coding of vergence from a spatial viewpoint by plotting disparity-tuning curves and from a temporal viewpoint by plotting the time course of the initial (open-loop) events. In both cases, the vergence information conveyed by the summed activity became apparent only after inverting the contributions of those cells whose tuning curves had negative slopes around zero disparity, which is roughly at the center of the servo range of the disparity-vergence system (Figs. 6 and 12). Given that the tuning curves for vergence had a positive slope around zero disparity, this inversion meant that all neurons would make a positive contribution to the population sum over the important servo range. Of course, for the brain to achieve a similar result would require appropriate excitatory and inhibitory connections. (Note that the sign of the disparity stimuli and the vergence responses was solely a matter of convention.)

SPATIAL ASPECTS. It is apparent from above that the single-cell data gave little hint that summing the disparity tuning curves for all of the units recorded from a given monkey would yield a curve strongly resembling the tuning curve for that animal's vergence response ($r^2 > 0.93$). This close match between summed activity and vergence was equally true for the data obtained with correlated and anticorrelated stimuli in spite of the fact that the associated vergence responses were very different. Further, the summed activity even reproduced the slight differences in the depth tuning of the vergence responses of the two monkeys from which we recorded most of our data (see Fig. 12, *A* and *B*). In fact, these idiosyncrasies were significant enough that the summed activity from one monkey fitted the vergence data for the other monkey only very poorly (r^2: 0.29, 0.35).

TEMPORAL ASPECTS. Our attempts to demonstrate the existence of temporal coding of vergence at the population level, by summing together all of the discharge profiles elicited by a given disparity step, were less successful. Visual inspection of the summed discharge profiles suggests that low signal to noise was a major factor here, emanating from irregularities in the spike trains coupled with inadequate sample sizes (Gomi et al. 1998). (Such noise factors were less evident in the analysis of spatial coding, presumably because it involved response measures that were averaged over time before summing.) Nonetheless, for 40% of the data sets obtained with correlated stimuli, the summed temporal profiles accounted for more than 90% of the stimulus-induced variation in the temporal profile of the associated vergence velocity. This is remarkable in that many of the unit response profiles had appreciable phasic components and showed little resemblance to the vergence response profiles: see, for example, the unit response profiles in Fig. 2, which formed part of a data set that when summed together fitted the associated vergence velocity profile with an r^2 value of 0.933. Clearly, the initial phasic components seen at the single-cell level largely disappeared at the population level, presumably because of the latency differences among the individual units (temporal summation).

COMPARISON WITH OTHER STUDIES. Our data suggest that the magnitude, direction, and time course of the initial vergence velocity responses associated with disparity steps are all encoded in the summed activity of the disparity-sensitive cells in MST. This brings to mind the vector addition model of motor cortex proposed by Georgopoulos and colleagues in which the direction of arm movement is reconstructed by summing the preferred direction vectors of the individual cells weighted by their firing rates (Georgopoulos et al. 1983, 1986, 1988; Kettner et al. 1988).

A population model that estimates the moving observer's direction of heading and incorporates MT and MST has been proposed by Lappe and Rauschecker (1993) and was recently used to successfully predict an optic flow illusion (Lappe and Duffy 1999). In this model, MST is laid out as a two-dimensional map of all the possible heading directions within the visual field, and each heading direction is represented by a separate population of neurons. The summed activity within each population peaks at a particular preferred heading and the direction of heading is signaled by the population showing the greatest activity, that is, by the location of the peak of activity on the map of heading directions. Thus the summation within MST in this model is local, whereas in our proposal, it is global.

Neural mediation of short-latency disparity-vergence eye movements

It is known that the initial vergence eye movements elicited by a step of disparity, which we have shown to be encoded in the summed activity of the disparity-sensitive cells in MST, are significantly reduced in amplitude by bilateral lesions in MST (Takemura et al. 1999, 2000). We suggest that these two observations are causally linked and that the vergence responses result at least in part from the population activity in MST. The genetic algorithm indicated that subsets of MST cells could reproduce the vergence responses better than our entire recorded population of cells, raising the possibility that only a subpopulation of MST cells contributes to the vergence responses. That the subsets of cells selected by the genetic algorithm always included cells from all four groups suggests that if only a subset of MST cells contributes to vergence, this subset is randomly selected from the entire population. There is strong evidence, based on lesions and single-unit recordings, that MST is critically involved in the generation of another type of short-latency eye movement, ocular following (OFR), elicited by large-field motion (see Kawano et al. 2000 for review) and the associated activity in MST precedes the motor responses by a time interval that is closely comparable with that found in the present study. Thus the estimated latency of the MST discharges preceded the estimated latency of the associated motor responses on average by 8.6 ms for OFR (Kawano et al. 1994) and by 8.9 ms for the vergence responses in the present study.

There are a number of putative pathways by which MST might produce vergence eye movements, including direct, subcortical projections. One of the latter is a projection to the dorsolateral pontine nuclei (Boussaoud et al. 1992; Glickstein et al. 1980, 1985), which have been shown to contain cells that discharge in relation to vergence (Zhang and Gamlin 1997) and project in turn to regions of the cerebellum (paraflocculus and vermis) that are known to be concerned with eye movements: for review, see Leigh and Zee (1999). There have been some preliminary reports that MST projects to the superior colliculus (Colby and Olson 1985; Lock et al. 1990), a structure that has recently been implicated in the production of vergence: see Chaturvedi and Van Gisbergen (2000) for recent review. Another direct subcortical projection from MST, to the nucleus of the optic tract, is thought to mediate OFR and optokinetic responses but, to date, is not known to have any involvement with vergence eye movements: see Inoue, Takemura, Kawano, and Mustari (2000) for recent review. MST is also known to project to two cortical areas that are interconnected and have been shown to contain neurons that discharge in relation to disparity stimuli and/or vergence eye movements: LIP (Gnadt and Mays 1995) and FEF (Ferraina et al. 2000; Gamlin et al. 1996). Neurons in LIP that carry depth-related information have been shown to project directly to the superior colliculus (Gnadt and Beyer 1998), and FEF neurons project directly to the medial part of the nucleus reticularis tegmenti pontis (Huerta et al. 1986; Leichnetz et al. 1984; Stanton et al. 1988), which shows vergence-related activity (Gamlin and Clarke 1995) and projects in turn to the premotor neurons for vergence (in the supraoculomotor and adjacent reticular formation) via the posterior interposed and fastigial nuclei of the cerebellum: see Gamlin, Yoon, and Zhang (1996) and Gamlin (1999) for review.

To the extent that the summed discharges of the disparity-selective neurons in MST carry a complete description of the associated vergence eye movements, they would have the potential to generate the entire vergence response, assuming appropriate dynamical processing in the projection pathways (Mays et al. 1986; Patel et al. 1997). The vergence deficits that follow lesions of MST are only partial (Takemura et al. 1999, 2000), leaving the possibility that other regions make a significant contribution (though it also seems very likely that the lesions were only partial). Of course, we do not know if the vergence system actually utilizes the population coding in MST, but were it to do so, then any attempt to correlate single-unit activity in MST with motor behavior would be a largely meaningless exercise.

Other data have implicated MST in perception, especially of self motion (Bradley et al. 1996; Britten and Wezel 1998; Duffy 1998; Duffy and Wurtz 1991a,b, 1995, 1997a,b; Graziano et al. 1994; Orban et al. 1995; Page and Duffy 1999; Roy et al. 1992; Saito et al. 1986; Tanaka et al. 1986, 1989). In our study, all of the MST neurons that responded to disparity steps applied to correlated patterns also responded when the same steps were applied to anticorrelated patterns even though the latter do not support depth perception and are seen as rivalrous (Cogan et al. 1993; Cumming and Parker 1997; Cumming et al. 1998; Masson et al. 1997). In fact, gradual changes in the disparity of large correlated patterns (similar to the ones we have used) have also been shown to produce little sensation of motion in depth (Erkelens and Collewijn 1985; Regan et al. 1986). This is in accord with the finding that stereopsis is much better for relative than for absolute disparity (Westheimer 1979), where absolute disparity is given by the difference in the locations of the two retinal images of a given object and relative disparity refers to the differences in the absolute disparity of different objects. Our stimuli contained only absolute disparity cues, and changes in the vergence angle between the

two eyes affect only the absolute disparity of the seen images. The implication is that the activity in MST that we have studied is linked to motor, rather than perceptual, processes.

APPENDIX A

Fuzzy clustering

The major advantage that fuzzy clustering algorithms have over classic (often referred to as hard or crisp) clustering algorithms is that they provide much more information about the data being clustered. More precisely, while crisp algorithms simply assign each data point to one out of a set of clusters, fuzzy algorithms compute, for each data point, a measure of how close that point is to the "center" of each cluster. This measure is termed the "fuzzy membership" of the point in a particular cluster: the higher the membership value, the closer the point to the center of that cluster. The choice of fuzzy clustering algorithms is very large (Hoppner et al. 1999), but we have decided to use the simplest, i.e., the fuzzy c-means clustering algorithm developed by Bezdek (1981), which assumes that the clusters are hyperspheres having approximately the same diameter. The concept behind the algorithm is simple. It minimizes the following functional

$$J_{\mathrm{m}}(\mathbf{U}, \mathbf{v}) = \sum_{k=1}^{n} \sum_{i=1}^{c} u_{ik}^{m} \cdot d_{ik}^{2} \tag{A1}$$

where c is the number of clusters, n is the number of data points, $\mathbf{U}$ is a matrix ($c \times n$) containing the (fuzzy) memberships of each point in each cluster ($U = [u_{ik}]$), $\mathbf{v}$ is a matrix of the cluster centers, m is a weighting exponent (usually called the "fuzzifier"), and d_{ik} is the Euclidean distance between the kth data point and the center of the ith cluster. The fuzzifier, m, can have any value >1, but we always used a value of 2 as it is the one most commonly used (Hoppner et al. 1999). The algorithm developed by Bezdek defines a set of rules that compute iteratively the cluster centers and the fuzzy memberships, until J_{m} reaches a minimum. The initial values of the cluster centers are chosen at random. The only constraints imposed by the algorithm are that the membership values can only vary between 0 and 1 and that the memberships of each individual point in the various clusters must add up to 1.

The most important parameter to be defined is then the number of clusters, c. Unfortunately, there is no golden rule for choosing it even though many so-called cluster validation measures have been proposed to solve this problem. These measures try to establish, in an objective way, which value of c is best for a given data set, but unfortunately they work well, and are in agreement with each other, only when the groups are fairly well separated, in which case a fuzzy algorithm is not necessary. Thus we decided to use the stability of the clustering solution (i.e., the convergence to the same solution regardless of the initial conditions) as an indication of the validity of the number of clusters selected, even though this is not a standard procedure. The detailed rationale behind our particular choice is indicated in RESULTS.

APPENDIX B

Genetic algorithm

In examining the correlation between single-unit activity in MST and disparity vergence, we attempted to identify the subset of cells whose aggregate behavior came closest to representing the vergence behavior. This was done by searching for the subset of cells whose disparity tuning curves, when summed together, best matched the disparity tuning curve for vergence. A mathematical formulation of the problem helps in deciding how to approach it. If VG_i is the vergence at different disparities ($i = 1 \ldots 11$) and UN_{ij} ($j = 1 \ldots n$) is the discharge rate of unit j for disparity i, for each subset of cells, $\mathbf{T}$, we can compute their summed tuning curve as

$$UN\{\mathbf{T}\} = \sum_{j \in T} UN_{ij} \tag{B1}$$

which can be considered as a generic solution. The task then is to find the optimal solution, that is the set, $\mathbf{T}$, such that $UN\{\mathbf{T}\}$ best fits the vergence behavior. Accordingly we searched for the set, $\mathbf{T}$, that minimizes the following mean square error

$$\mathrm{MSE}\ \{\mathbf{T}\} = \frac{\sum_{i=1}^{11} [UN\{\mathbf{T}\}_i - (g \cdot VG_i + b)]^2}{11} \tag{B2}$$

with b (bias) and g (gain) chosen differently every time to get the lowest possible MSE for a given $\mathbf{T}$. (As this is a very simple minimization problem, optimal values for b and g can be found using the Nelder-Mead simplex algorithm.) We can formulate the problem differently by defining a vector, S_T, having n elements (1 for each unit) such that the ith element of S_T is 1 if the ith unit belongs to $\mathbf{T}$ and is 0 otherwise. As there is a one-to-one relationship between the vector $\mathbf{T}$ and the string of bits S_T, we can associate each S_T with its corresponding summed tuning curve ($UN\{S_T\} = UN\{T\}$) and mean square error ($MSE\{S_T\} = MSE\{T\}$). Clearly, there are 2^n different vectors S_T (i.e., potential solutions). In the present study, we obtained disparity tuning curves for 49 units from one animal, hence for that animal there are 2^{49} potential solutions. As the complexity of the problem ruled out the use of brute force, we had to resort to some minimization technique and decided to use a genetic algorithm (GA). From an operational standpoint, GAs applied to problems in which the variables that define a solution are binary are straightforward. They start with an initial set of *chromosomes*, where each chromosome represents a set of values of the binary variables (in our case, a vector S_T) and thus can be associated with a solution of the problem (in our case $UN\{S_T\}$). Each chromosome consists of a string of n genes, where each gene represents a binary variable (in our case, each gene is associated with 1 of the disparity tuning curves obtained for the n units recorded from a given animal) that can have a value of either 1, indicating that the gene is *expressed* (in our case, that the corresponding tuning curve is included in the sum $UN\{S_T\}$), or 0, indicating that the gene *is not expressed* (in our case, not included in $UN\{S_T\}$). For the first *generation*, each gene is assigned a value of 0 or 1 with equal probability. Pairs of chromosomes are then combined, according to a defined set of rules, to find a new generation of chromosomes. The *mating rules* are such that they tend to apply some *evolutionary pressure* so that chromosomes that are associated with better solutions of the problem have a higher probability of mating (the assumption being that better parents generate better offspring). Thus GAs require a measure of how good a given solution is; this measure is usually referred to as the *fitness* of the solution. Of course, this measure is problem-specific. In our case, each chromosome represents an S_T, and the solution to the problem is represented by the sum $UN\{S_T\}$ of all the disparity tuning curves associated with genes that are expressed (i.e., that have a value of 1) in that particular chromosome. The fitness of this solution is then represented by how well this sum fits the disparity tuning curve for the vergence response. Accordingly, the $MSE\{S_T\}$ represents a direct measure of the fitness of the solution (and thus of the chromosome): the lower the MSE, the fitter the chromosome. The algorithm will then find chromosomes (i.e., sets S_T) that have higher and higher fitness (i.e., lower and lower $MSE\{S_T\}$) at each generation. We used 5,000 vectors/chromosomes, and the following standard evolutionary rules (e.g., see Goldberg 1989; Mitchell 1998) were then used to create new generations, each having 5,000 individual vectors/chromosomes: *1*) Five percent of the chromosomes, randomly selected, were passed on to the next generation unchanged. *2*) Five percent of the chromosomes, randomly selected, were passed

on to the next generation after changing the values of 1–3 genes randomly (*mutations*). *3*) The single chromosome having the lowest overall MSE was passed on to the next generation unchanged (*elitarism*). *4*) The remaining chromosomes (almost 90%) were each mated with the best (i.e., the chromosome with the smallest MSE) of 10 randomly selected chromosomes; the mating process involved selecting each of the *n* genes at random from either of the parents.

Rules 1 and 2 keep some diversity in the population, whereas rules 3 and 4 exert evolutionary pressure. The algorithm ran for 50 generations, during which the scatter, the mean and the minimum MSE for the population gradually diminished, usually stabilizing after ~30 generations. After 50 generations, the vast majority of the chromosomes were identical. This surviving chromosome represents the algorithm's best estimate of the optimal solution, i.e., the subset of units whose summed disparity tuning curves best correlates with the disparity tuning curve for vergence. Following standard practice, we ran the algorithm numerous (100) times for each data set, both to validate our finding and to give the algorithm the opportunity to sample as many regions of the solution space as possible. Of course this does not allow us to conclude that we have found the optimal solution (as that would require testing all the possible solutions), but it is the best estimate given the technologies available at this time.

We thank Drs. L. M. Optican, Y. Kodaka, M. Shidara, and S. Yamane for valuable advice; M. Okui, A. Kameyama, and T. Takasu for technical assistance; and Y. Yaguchi and S. Inoue for secretarial help.

This research was supported by grants from the Japanese Agency of Industrial Science and Technology, Core Research for Evolutional Science and Technology (CREST) of Japan Science and Technology Corporation, the Human Frontier Science Program, Japan Society for the Promotion of Science Research Fellowships for Young Scientists, and the National Eye Institute.

REFERENCES

ANDERSEN RA, ASANUMA C, ESSICK G, AND SIEGEL RM. Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J Comp Neurol* 296: 65–113, 1990.

BEZDEK JC. *Pattern Recognition with Fuzzy Objective Function Algorithms*. New York: Plenum, 1981.

BLATT GJ, ANDERSEN RA, AND STONER GR. Visual receptive field organization and cortico-cortical connections of the lateral intraparietal area (area LIP) in the macaque. *J Comp Neurol* 299: 421–445, 1990.

BOUSSAOUD D, DESIMONE R, AND UNGERLEIDER LG. Subcortical connections of visual areas MST and FST in macaques. *Vis Neurosci* 9: 291–302, 1992.

BRADLEY DC AND ANDERSEN RA. Center-surround antagonism based on disparity in primate area MT. *J Neurosci* 18: 7552–7565, 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.

BRADLEY DC, QIAN N, AND ANDERSEN RA. Integration of motion and stereopsis in middle temporal cortical areas of macaques. *Nature* 373: 609–611, 1995.

BRITTEN KH AND VAN WEZEL RJ. Electrical microstimulation of cortical area MST biases heading perception in monkeys. *Nat Neurosci* 1: 59–63, 1998.

BURKHALTER A AND VAN ESSEN DC. Processing of color, form and disparity information in visual areas VP and V2 of ventral extrastriate cortex in the macaque monkey. *J Neurosci* 6: 2327–2351, 1986.

BUSETTINI C, MILES FA, AND KRAUZLIS RJ. Short-latency disparity vergence responses and their dependence on a prior saccadic eye movement. *J Neurophysiol* 75: 1392–1410, 1996.

CHATURVEDI V AND VAN GISBERGEN JAM. Stimulation in the rostral pole of monkey superior colliculus: effects on vergence eye movements. *Exp Brain Res* 132: 72–78, 2000.

COGAN AI, LOMAKIN AJ, AND ROSSI AF. Depth in anticorrelated stereograms: effects of spatial density and interocular delay. *Vision Res* 33: 1959–1975, 1993.

COLBY CL AND OLSON CR. Visual topography of cortical projections to monkey superior colliculus. *Soc Neurosci Abstr* 11: 1244, 1985.

CRIST CF, YAMASAKI DSG, KOMATSU H, AND WURTZ RH. A grid system and a microsyringe for single cell recording. *J Neurosci Methods* 26: 117–122, 1988.

CUMMING BG AND JUDGE SJ. Disparity-induced and blur-induced convergence eye movement and accommodation in the monkey. *J Neurophysiol* 55: 896–914, 1986.

CUMMING BG AND PARKER AJ. Responses of primary visual cortical neurons to binocular disparity without depth perception. *Nature* 389: 280–283, 1997.

CUMMING BG AND PARKER AJ. Binocular neurons in V1 of awake monkeys are selective for absolute, not relative, disparity. *J Neurosci* 19: 5602–5618, 1999.

CUMMING BG AND PARKER AJ. Local disparity not perceived depth is signaled by binocular neurons in cortical area V1 of the macaque. *J Neurosci* 15: 4758–4767, 2000.

CUMMING BG, SHAPIRO SE, AND PARKER AJ. Disparity detection in anticorrelated stereograms. *Perception* 27: 1367–1377, 1998.

DEANGELIS GC, CUMMING BG, AND NEWSOME WT. Cortical area MT and the perception of stereoscopic depth. *Nature* 394: 677–680, 1998.

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

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

DUFFY CJ AND 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, 1991a.

DUFFY CJ AND WURTZ RH. Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. *J Neurophysiol* 65: 1346–1359, 1991b.

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. Medial superior temporal neurons respond to speed patterns in optic flow. *J Neurosci* 17: 2839–2851, 1997a.

DUFFY CJ AND WURTZ RH. Planar directional contributions to optic flow responses in MST neurons. *J Neurophysiol* 77: 782–796, 1997b.

EFRON B AND TIBSHIRANI R. Statistical data analysis in the computer age. *Science* 253: 390–395, 1991.

EIFUKU S AND WURTZ RH. Response to motion in extrastriate area MSTl: disparity sensitivity. *J Neurophysiol* 82: 2462–2475, 1999.

ERKELENS CJ AND COLLEWIJN H. Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Res* 25: 583–588, 1985.

FELLEMAN DJ AND VAN ESSEN DC. Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *J Neurophysiol* 57: 889–920, 1987.

FERRAINA S, PARÉ M, AND WURTZ RH. Disparity sensitivity of frontal eye field neurons. *J Neurophysiol* 83: 625–629, 2000.

GALLYAS F. Silver staining of myelin by means of physical development. *Neurol Res* 1: 203–209, 1979.

GAMLIN PD. Subcortical neural circuits for ocular accommodation and vergence in primates. *Ophthalmol Physiol Opt* 19: 81–89, 1999.

GAMLIN PD AND CLARKE RJ. Single-unit activity in the primate nucleus reticularis tegmenti pontis related to vergence and ocular accommodation. *J Neurophysiol* 73: 2115–2119, 1995.

GAMLIN PD, YOON K, AND ZHANG H. The role of the cerebro-ponto-cerebellar pathways in the control of vergence eye movements. *Eye* 10: 167–171, 1996.

GEORGOPOULOS AP, CAMINITI R, KALASKA JF, AND MASSEY JT. Spatial coding of movement: a hypothesis concerning the coding of movement direction by motor cortical populations. *Exp Brain Res Suppl* 7: 327–336, 1983.

GEORGOPOULOS AP, KETTNER RE, AND SCHWARTZ AB. Primate motor cortex and free arm movements to visual targets in three-dimensional space. II. Coding of the direction of movement by a neuronal population. *J Neurosci* 8: 2928–2937, 1988.

GEORGOPOULOS AP, SCHWARTZ AB, AND KETTNER RE. Neuronal population coding of movement direction. *Science* 233: 1416–1419, 1986.

GLICKSTEIN M, COHEN JL, DIXON B, GIBSON A, HOLLINS M, LABOSSIERE E, AND ROBINSON F. Corticopontine visual projections in macaque monkeys. *J Comp Neurol* 190: 209–229, 1980.

GLICKSTEIN M, MAY JG, AND MERCIER BE. Corticopontine projection in the macaque: the distribution of labelled cortical cells after large injections of horseradish peroxidase in the pontine nuclei. *J Comp Neurol* 235: 343–359, 1985.

GNADT JW AND BEYER J. Eye movements in depth: what does the monkey's parietal cortex tell the superior colliculus? *Neuroreport* 9: 233–238, 1998.

GNADT JW AND MAYS LE. Neurons in monkey parietal area LIP are tuned for eye movement parameters in three-dimensional space. *J Neurophysiol* 73: 280–297, 1995.

GOLDBERG DE. *Genetic Algorithms in Search. Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.

GOMI H, SHIDARA M, TAKEMURA A, INOUE Y, KAWANO K, AND KAWATO M. Temporal firing patterns of Purkinje cells in the cerebellar ventral paraflocculus during ocular following responses in monkeys. I. Simple spikes. *J Neurophysiol* 80: 818–831, 1998.

GRAZIANO M, ANDERSEN R, AND SNOWDEN R. Tuning of MST neurons to spiral motions. *J Neurosci* 14: 54–56, 1994.

HOPPNER F, KLAWONN F, KRUSE R, AND RUNKLER T. *Fuzzy Cluster Analysis*. Chichester: Wiley, 1999.

HUBEL DH AND LIVINGSTONE MS. Segregation of form, color and stereopsis in primate area 18. *J Neurosci* 7: 3378–3415, 1987.

HUBEL DH AND WIESEL TN. Cells sensitive to binocular depth in area 18 of the macaque monkey cortex. *Nature* 225: 41–42, 1970.

HUERTA MF, KRUBITZER LA, AND KAAS JH. Frontal eye field as defined by intracortical microstimulation in squirrel monkeys, owl monkeys, and macaque monkeys. I. Subcortical connections. *J Comp Neurol* 253: 415–439, 1986.

INOUE Y, TAKEMURA A, KAWANO K, AND MUSTARI MJ. Role of the pretectal nucleus of the optic tract in short-latency ocular following responses in monkeys. *Exp Brain Res* 131: 269–281, 2000.

JONES R. Fusional vergence: sustained and transient components. *Am J Optom Physiol Optics* 57: 640–644, 1980.

JUDGE SJ, RICHMOND BJ, AND CHU FC. Implantation of magnetic search coils for measurement of eye position: an improved method. *Vision Res* 20: 535–538, 1980.

KAWANO K, INOUE Y, TAKEMURA A, KODAKA Y, AND MILES FA. The role of MST neurons during ocular tracking in 3-D space. In: *Neuronal Processing of Optic Flow*, edited by Lappe M. New York: Academic, 2000, p. 49–63.

KAWANO K, SHIDARA M, WATANABE Y, AND YAMANE S. Neural activity in cortical area MST of alert monkey during ocular following responses. *J Neurophysiol* 71: 2305–2324, 1994.

KAWANO K, SHIDARA M, AND YAMANE S. Neural activity in dorsolateral pontine nucleus of alert monkey during ocular following responses. *J Neurophysiol* 67: 680–703, 1992.

KETTNER RE, SCHWARTZ AB, AND GEORGOPOULOS AP. Primate motor cortex and free arm movements to visual targets in three-dimensional space. III. Positional gradients and population coding of movement direction from various movement origins. *J Neurosci* 8: 2938–2947, 1988.

LAPPE M AND DUFFY CJ. Optic flow illusion and single neuron behaviour reconciled by a population model. *Eur J Neurosci* 11: 2323–2331, 1999.

LAPPE M AND RAUSCHECKER JP. A neural network for the processing of optic flow from ego-motion in man and higher mammals. *Neural Comput* 5: 374–391, 1993.

LEICHNETZ G, SMITH DJ, AND SPENCER RF. Cortical projections to the paramedian tegmental and basilar pons in the monkey. *J Comp Neurol* 228: 388–408, 1984.

LEIGH RJ AND ZEE DS. *The Neurology of Eye Movements* (3rd ed.). London: Oxford Univ. Press, 1999.

LOCK TM, BAIZER JS, AND BENDER DB. Distribution of corticotectal cells in the superior temporal sulcus of the macaque. *Soc Neurosci Abstr* 16: 110, 1990.

MACPHERSON JM AND ALDRIDGE JW. A quantitative method of computer analysis of spike train data collected from behaving animals. *Brain Res* 175: 183–187, 1979.

MASSON GS, BUSETTINI C, AND MILES FA. Vergence eye movements in response to binocular disparity without depth perception. *Nature* 389: 283–286, 1997.

MAUNSELL JHR AND VAN ESSEN DC. Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. *J Neurophysiol* 49: 1148–1167, 1983a.

MAUNSELL JHR AND VAN ESSEN DC. The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. *J Neurosci* 3: 2563–2586, 1983b.

MAYS LE, PORTER JD, GAMLIN PDR, AND TELLO CA. Neural control of vergence eye movements: neurons encoding vergence velocity. *J Neurophysiol* 56: 1007–1021, 1986.

MITCHELL DE. Properties of stimuli eliciting vergence eye movements and stereopsis. *Vision Res* 10: 145–162, 1970.

MITCHELL M. *An Introduction to Genetic Algorithms*. Cambridge, MA: MIT Press, 1998.

NEWSOME WT, WURTZ RH, AND KOMATSU H. Relation of cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. *J Neurophysiol* 60: 604–620, 1988.

NOMURA M, MATSUMOTO G, AND FUJIWARA S. A binocular model for the simple cell. *Biol Cybern* 63: 237–242, 1990.

OHZAWA I. Mechanisms of stereoscopic vision: the disparity energy model. *Curr Opin Neurobiol* 8: 509–515, 1998.

OHZAWA I, DEANGELIS GC, AND FREEMAN RD. Stereoscopic depth discrimination in the visual cortex: neurons ideally suited as disparity detectors. *Science* 249: 1037–1041, 1990.

ORBAN GA, LAGAE L, RAIGUEL S, XIAO D, AND MAES H. The speed tuning of middle superior temporal (MST) cell responses to optic flow components. *Perception* 24: 269–285, 1995.

PAGE WK AND DUFFY CJ. MST neuronal responses to heading direction during pursuit eye movements. *J Neurophysiol* 81: 596–610, 1999.

PATEL SS, OGMEN H, WHITE JM, AND JIANG BC. Neural network model of short-term horizontal disparity vergence dynamics. *Vision Res* 37: 1383–1399, 1997.

POGGIO GF AND FISCHER B. Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. *J Neurophysiol* 40: 1392–1405, 1977.

POGGIO GF, GONZALEZ F, AND KRAUSE F. Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. *J Neurosci* 8: 4531–4550, 1988.

POGGIO GF AND TALBOT WH. Mechanisms of static and dynamic stereopsis in foveal cortex of the rhesus monkey. *J Physiol (Lond)* 315: 469–492, 1981.

PRINCE SJ, POINTON AD, CUMMING BG, AND PARKER AJ. The precision of single neuron responses in cortical area V1 during stereoscopic depth judgments. *J Neurosci* 20: 3387–3400, 2000.

RASHBASS C AND WESTHEIMER G. Disjunctive eye movements. *J Physiol (Lond)* 159: 339–360, 1961.

REGAN D, ERKELENS CJ, AND COLLEWIJN H. Necessary conditions for the perception of motion in depth. *Invest Ophthalmol Vis Sci* 27: 584–597, 1986.

RICHMOND BJ, OPTICAN LM, PODELL M, AND SPITZER H. Temporal encoding of two-dimensional patterns by single units in primate inferior temporal cortex. I. Response characteristics. *J Neurophysiol* 57: 132–146, 1987.

ROY J-P, KOMATSU H, AND WURTZ RH. Disparity sensitivity of neurons in monkey extrastriate area MST. *J Neurosci* 12: 2478–2492, 1992.

SAITO H, YUKIE M, TANAKA K, HIKOSAKA K, FUKADA Y, AND IWAI E. Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J Neurosci* 6: 145–157, 1986.

SAKATA H, SHIBUTANI H, AND KAWANO K. Functional properties of visual tracking neurons in posterior parietal association cortex of the monkey. *J Neurophysiol* 49: 1364–1380, 1983.

SCHALL JD, MOREL A, KING DJ, AND BULLIER J. Topography of visual cortex connections with frontal eye field in macaque: convergence and segregation of processing streams. *J Neurosci* 15: 4464–4487, 1995.

SHENOY KV, BRADLEY DC, AND ANDERSEN RA. Influence of gaze rotation on the visual response of primate MSTd neurons. *J Neurophysiol* 81: 2764–2786, 1999.

SMITH EL III, CHINO Y, NI J, AND CHENG H. Binocular combination of contrast signals by striate cortical neurons in the monkey. *J Neurophysiol* 78: 366–382, 1997.

SQUATRITO S AND MAIOLI MG. Gaze field properties of eye position neurones in areas MST and 7a of the macaque monkey. *Vis Neurosci* 13: 385–398, 1996.

SQUATRITO S AND MAIOLI MG. Encoding of smooth pursuit direction and eye position by neurons of area MSTd of macaque monkey. *J Neurosci* 17: 3847–3860, 1997.

STANTON GB, BRUCE CJ, AND GOLDBERG ME. Frontal eye field efferents in the macaque monkey. II. Topography of terminal fields in midbrain and pons. *J Comp Neurol* 271: 493–506, 1988.

TAKEMURA A, INOUE Y, AND KAWANO K. The role of MST neurons in short-latency visual tracking eye movements. *Soc Neurosci Abstr* 26: 1715, 2000.

TAKEMURA A, INOUE Y, KAWANO K, AND MILES FA. Neuronal activity in medial superior temporal area of alert monkeys associated with ultra-short latency disparity vergence. *Forum Eur Neurosci Abstr* 10: 190, 1998.

TAKEMURA A, INOUE Y, KAWANO K, QUAIA C, AND MILES FA. Short-latency discharges in medial superior temporal area of alert monkey to sudden changes in the horizontal disparity. *Soc Neurosci Abstr* 23: 1557, 1997.

TAKEMURA A, INOUE Y, KAWANO K, QUAIA C, AND MILES FA. Evidence that disparity-sensitive cells in medial superior temporal area contribute to

short-latency vergence eye movements. *Soc Neurosci Abstr* 25: 1400, 1999.

TANAKA K, FUKADA Y, AND SAITO H. Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *J Neurophysiol* 62: 642–656, 1989.

TANAKA K, HIKOSAKA K, SAITO H, YUKIE M, FUKADA Y, AND IWAI E. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci* 6: 134–144, 1986.

THIER P AND ERICKSON RG. Responses of visual-tracking neurons from cortical area MST-l to visual, eye and head motion. *Eur J Neurosci* 4: 539–553, 1992.

TROTTER Y, CELEBRINI S, STRICANNE B, THORPE S, AND IMBERT M. Neural processing of stereopsis as a function of viewing distance in primate visual cortical area VI. *J Neurophysiol* 76: 2872–2885, 1996.

UNGERLEIDER LG AND DESIMONE R. Cortical connections of visual area MT in the macaque. *J Comp Neurol* 248: 190–222, 1986.

WESTHEIMER G. Cooperative neural processes involved in stereoscopic acuity. *Exp Brain Res* 36: 585–597, 1979.

WESTHEIMER G AND MITCHELL AM. Eye movement responses to convergent stimuli. *Arch Ophthalmol* 55: 848–856, 1956.

WURTZ RH. Visual receptive fields of striate cortex neurons in awake monkeys. *J Neurophysiol* 32: 727–742, 1969.

ZHANG HY AND GAMLIN PDR. The dorsolateral pontine nucleus of the primate: neurons related to vergence and accommodation. *Soc Neurosci Abstr* 23: 1557, 1997.