The role of the posterior parietal cortex in coordinate transformations for visual–motor integration[1]

RICHARD A. ANDERSEN[2]

Department of Brain and Cognitive Sciences, E25-236, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139, U.S.A.

AND

DAVID ZIPSER

Institute for Cognitive Science, University of California San Diego, La Jolla, CA, U.S.A.

Received November 16, 1987

ANDERSEN, R. A., and ZIPSER, D. 1988. The role of the posterior parietal cortex in coordinate transformations for visual–motor integration. Can. J. Physiol. Pharmacol. **66**: 488–501.

Lesion to the posterior parietal cortex in monkeys and humans produces spatial deficits in movement and perception. In recording experiments from area 7a, a cortical subdivision in the posterior parietal cortex in monkeys, we have found neurons whose responses are a function of both the retinal location of visual stimuli and the position of the eyes in the orbits. By combining these signals area 7a neurons code the location of visual stimuli with respect to the head. However, these cells respond over only limited ranges of eye positions (eye-position-dependent coding). To code location in craniotopic space at all eye positions (eye-position-independent coding) an additional step in neural processing is required that uses information distributed across populations of area 7a neurons. We describe here a neural network model, based on back-propagation learning, that both demonstrates how spatial location could be derived from the population response of area 7a neurons and accurately accounts for the observed response properties of these neurons.

ANDERSEN, R. A., et ZIPSER, D. 1988. The role of the posterior parietal cortex in coordinate transformations for visual–motor integration. Can. J. Physiol. Pharmacol. **66**: 488–501.

Une lésion du cortex pariétal postérieur chez les singes et les humains provoque des carences spatiales dans le mouvement et la perception. En enregistrant dans la zone 7a chez des singes, une sous-division du cortex pariétal postérieur, nous avons trouvé des neurones dont les réponses sont fonction de la localisation rétinienne des stimuli visuels et de la position des yeux dans les orbites. En combinant ces signaux, les neurones de la zone 7a codent la localisation des stimuli visuels par rapport à la tête. Toutefois, ces cellules ne répondent que dans des gammes limitées de positions oculaires (codage fonction de la position des yeux). Le codage de la localisation dans l'espace craniotopique à toutes les positions oculaires (codage indépendant de la position des yeux) nécessite une étape additionnelle dans le processus neuronal qui utilise l'information distribuée dans les populations de neurones de la zone 7a, Ici, nous décrivons un modèle de réseau neuronal, basé sur la rétro-propagation de l'apprentissage, qui démontre comment la localisation spatiale pourrait dériver de la réponse de la population de neurones de la zone 7a et qui explique précisément les propriétés des réponses de ces neurones.

[Traduit par la revue]

Introduction

Neurophysiological experiments can derive very accurate information about the activity of single cells in relation to behavior. However, an understanding of brain function requires that the activity of these single cells be understood in the context of the entire network of nerve cells of which they are a member. Thus recent work on the behavior of networks with brain-like architecture is beginning to provide tools for the potential understanding of real brain circuits. In general, these network models are still very far removed from the structure of actual brains and as a result cannot provide literal models of the brain; however, they can begin to provide at a more abstract level general ideas about the types of coding that can occur in parallel networks such as that seen in the brain and how these codes might be represented at the single cell level.

In the study described in this paper combined neurophysiological and computational approaches were used to investigate how the brain represents visual space in craniotopic coordinates. The area of the brain implicated from the lesion literature to contain such a representation is the posterior parietal cortex (see Andersen 1987 for review). Our recording experiments in this area have indicated an unexpected neural code for craniotopic space that was only unambiguous in the population response. In the computational studies to be described a computer-simulated network model was used to learn stimulus location in craniotopic space based on eye and retinal position inputs modeled on the same types of inputs that do in fact impinge on the posterior parietal cortex. Interestingly the internal or "hidden" layer units in the network that are between the input and output units develop the same distributed coding that is found in the brain.

Spatial representations

There are several lines of evidence that indicate the brain uses non-retinotopic representations of space. For instance, one can reach accurately to the location of visual targets without visual feedback and independent of eye position, head position, or the location of the image on the retinas. Thus, the motor system must be use representations of visual stimuli mapped in body-centered coordinates rather than retinal coordinates. There is substantial evidence that during the planning of eye movements there is a stage in which the target of the eye movement is represented in craniotopic coordinates (Robinson 1975; Hallet and Lightstone 1976; Mays and Sparks 1980). Finally, we know by introspection that the visual world appears perceptually stable in spite of the fact that we are constantly making eye movements and subsequently shifting the location of images on the retinas. These results suggest that information

[1]This paper was presented at the IX International Symposium of the Centre de recherche en sciences neurologiques, Université de Montréal, entitled Spatial Representations and Sensorimotor Transformations (Montréal, Qué., May 28–29, 1987), and has undergone the Journal's usual peer review.

[2]Author to whom correspondence should be sent.

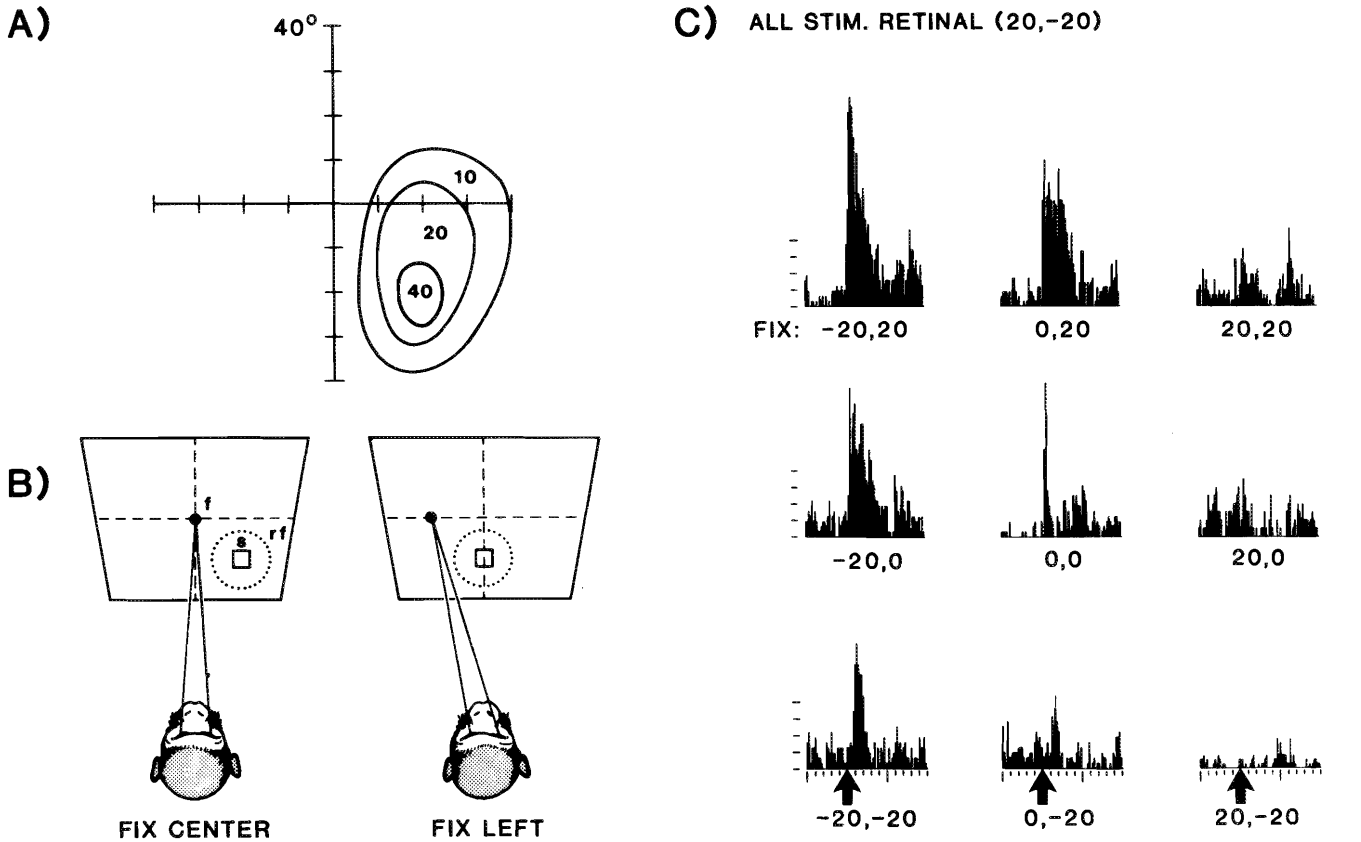

FIG. 1. (A) Receptive field of a neuron plotted in coordinates of visual angle determined with the animal always fixating straight ahead (screen coordinates 0,0). The contours represent the mean increased response rates in spikes per second. (B) Method of determining spatial gain fields of area 7a neurons. The animal fixates point f at different locations on the screen with his head fixed. The stimulus, s, is always presented in the center of the receptive field, rf. (C) Spatial gain field of the cell in (A). The poststimulus histograms are positioned to correspond to the locations of the fixations on the screen at which the responses were recorded for retinotopically identical stimuli presented in the center of the receptive field (histogram ordinate, 25 spikes per division, and abscissa, 100 ms per division; arrows indicate onset of stimulus flash). From Andersen et al. (1985), reproduced from Science (Washington, DC), 230: 456–458, © AAAS.

about eye position is used to compensate for movements of the visual scene. Another supporting observation comes from the fact that when the eyes are moved passively (by applying external pressure to the eyeball), the same retinal stimulus occurs as that seen in a willed movement, but in this case the world does indeed appear to move.

Location of spatial representations

The most likely area of the brain in which to find nonretinotopic representations of visual space is the posterior parietal cortex. Lesions to this area in humans and monkeys produce visual disorientation, a syndrome in which the subjects cannot reach accurately to visual targets and have difficulty navigating around seen obstacles (see Andersen 1987 for review). The patients are not blind and when tested often have normal visual field functions. However, they appear to be unable to associate what they see with the positions of their bodies.

Recording experiments

We decided to examine the coordinate frame for visual space represented in the posterior parietal cortex by mapping visual receptive fields with animals looking in different directions. The animals' heads were fixed to simplify the coordinate space to a head-centered frame. We reasoned that if the receptive fields moved with the eyes then the coordinate frame was retinotopic, but if they remained static in space then they were coding in at least craniotopic coordinates. Figure 1 shows how this experiment is done. The receptive field is first mapped with a flashed stimulus while the animal fixates a small fixation spot located straight ahead (0,0 in screen coordinates). Figure 1A shows a typical receptive field mapped in this way where the axes represent screen coordinates and the contour lines represent different levels of neural response. This receptive field is located down and to the right, is 30–40° in diameter, and gives an excitatory response to the flashed stimulus throughout its extent. Once the receptive field had been mapped, we then presented the stimulus at the retinotopic location that gave the maximum response, but with the animal gazing in different directions. If the response changes with eye position for retinotopically identical stimuli, then the cells are coding in something other than simple retinotopic coordinates. Figure 1B shows how the visual response is tested at different eye positions. The dotted line "rf" represents the receptive field that was mapped in Fig. 1A. Stimulus "s" is flashed at the most responsive location in the receptive field, in this case it is located in approximately the center of the response zone. On the left the animal is first required to fixate straight ahead (0,0 in screen coordinates). On the right the animal has been required to fixate to the left of straight ahead by 20° (−20,0 in screen coordinates). Since the head is fixed, the eyes are now in different positions in the orbits. Again the target is flashed in the same retinotopic location; however, since the eyes have moved 20° to the left, the

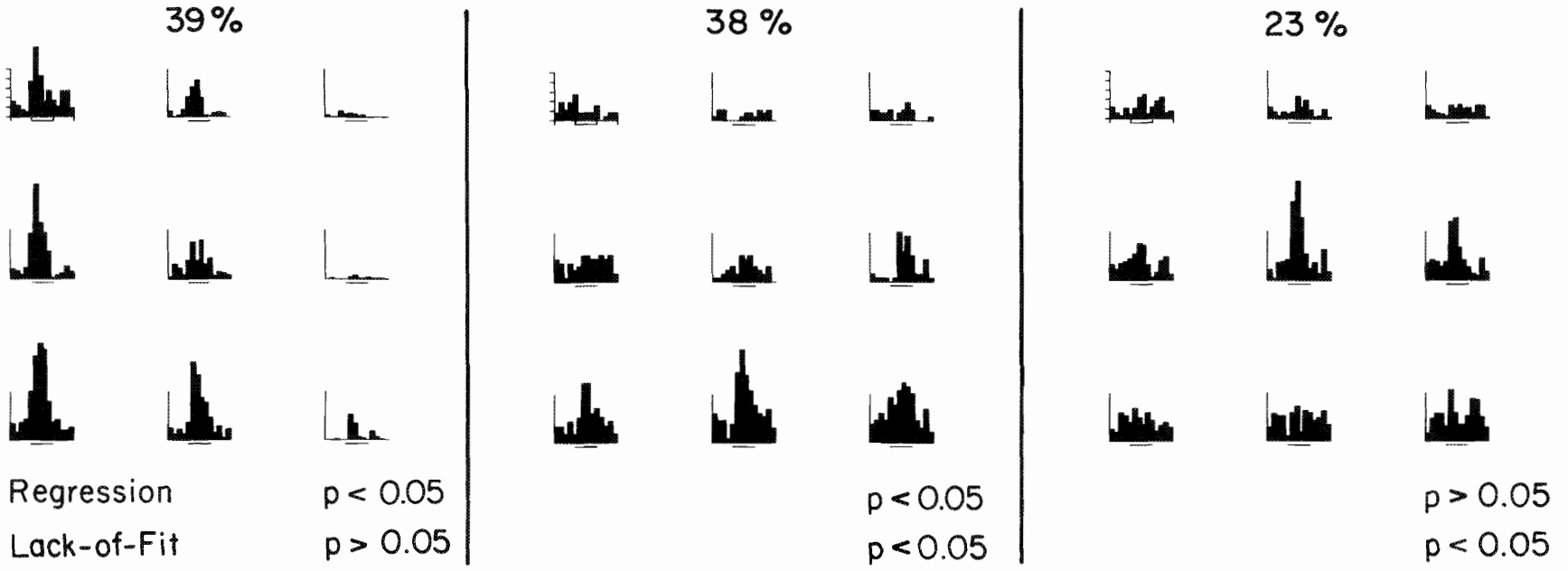

FIG. 2. Examples of gain fields analysed with a first order linear model with horizontal and vertical eye positions as independent variables. Each spatial gain field was generated by collecting eight repetitions of retinotopically identical stimuli delivered at each of nine eye positions using a randomized block design. The visually elicited response rates were adjusted for background activity by subtracting the rate of activity during the presentation of the visual stimulus from the rate of activity recorded before stimulus presentation at the same eye position. Conventional linear regression techniques were applied to partition the response variability into components dependent on x and y eye positions and residual "pure error" and "lack-of-fit" components for statistical testing. For the evoked visual responses, 39% showed a significant fit to the model and no significant lack of fit. An example of one of these cells is shown in the left panel. For another 38%, there was a significant planar component but also a significant lack of fit. These cells were often largely planar but had a small bump or depression in the plane. An example of one of these cells is shown in the middle panel; a large response for the middle histogram of the lower row of this gain field resulted in a significant lack of fit. Finally, 23% of the neurons had no significant planar component and a significant lack of fit. An example of one of these cells is shown in the rightmost panel. If one considers those cells of the middle category that were judged to be mostly planar, then 55% of the evoked response gain fields were planar. On the other hand, when the same cells were analysed for the total response (i.e., without adjusting for background activity), then 78% of the cells were found to have planar or nearly planar gain fields.

stimulus has also been moved to the left on the screen 20° such that the stimulus falls on the same retinotopic location.

Nine eye positions are tested in this manner. Results from this cell are shown in Fig. 1C. Each histogram is plotted at the corresponding fixation location. The cell was active for fixations up 20° and left 20°, less active for looking straight ahead, and not active at all when looking down 20° and right 20°. Thus the activity of the cell for retinotopically identical stimuli varied as a function of the angle of gaze. We refer to the plots for the nine fixation positions as shown in Fig. 1C as spatial gain fields. Notice that this particular gain field can be described by a plane tilted up and to the left.

We recorded complete spatial gain fields for 87 neurons. The mean evoked responses of these plots were further analysed using a first order linear model with independent variables of horizontal and vertical eye position to determine how many of these gain fields could be fit to a plane. The evoked activity was computed by subtracting the background activity before the test stimulus from the overall activity, during, and just after the stimulus. Sixteen cells were judged to have no gain fields for the evoked activity and were not analysed further. Figure 2 shows three categories of gain fields that were obtained by this analysis. Thirty-nine percent of the gain fields showed a significant planar component and no significant lack of fit indicating that a plane was the best model for the data. The example in the left panel could be fit with a plane that was tilted up when the animal fixated to the left and a small tilt component was also present for downward fixation. Another 38% showed a planar component but also a significant lack of fit, indicating that although a plane could be fit to the data, a plane was not the optimal model. The gain field in the middle panel shows a typical example for this group; usually there was a bump, in this case the lower middle histogram, in the overall planar fit, in this case a plane tilted upward for downward fixations. About one-half of these neurons looked very planar. Finally, another 23% showed no planar component and a significant lack of fit. The right panel shows a typical gain field from this group, which usually had peaked gain fields. Looking at these data as a whole, it is important to note that a majority of cells showed planar or largely planar gain fields (55%) but that there is also a significant number of gain fields (45%) that are not planar. Interestingly, when the same cells were analysed for their overall activity rather than just the evoked activity, it was found that a large proportion of the cells (78%) had planar or largely planar gain fields. Thus the total signal of background activity and evoked activity shows a greater degree of planarity than just the evoked response, and it is the most likely signal used by the brain for spatial localization, since it is the output of the neurons.

Controls

A major concern is that the visual background, which is imaged at different locations on the retinas at different angles of gaze, is influencing the responsiveness of the cells to the test flash. Two controls were performed to eliminate this possibility. Many of the recordings were made in complete darkness except for the stimulus and the fixation point so that there was no background. In fact, all of the gain fields in Figs. 1 and 2 were collected under these darkness conditions.

The second control was to change the angle of gaze using prisms; this requires the animal to change eye position without changing the retinal locations of the imaged background. Figure 3 shows an example of recordings using this prism control. The upper two histograms show responses to retinotopically identical stimuli with the animal looking left, which gave a big response, and looking right, which gave very little response. The animal was made to look left and right by moving the fixation target on the screen in a lighted room. The lower two histograms show responses to the same retinotopic stimuli, but now the fixation target is not moved on the screen and the animal is made to look left and right by viewing through base right and base left prisms, respectively. The eye movement recordings show that the animal is fixating at the same gaze angles as in the "no prisms" condition. In this case the retinotopic locations of the background image are the same and again the cell responds best to the stimulus with leftward fixation. Thus we can conclude that the effect of eye position on visual responses for this cell was not a result of background shifts. We found the same results in 11 out of 12 cells tested in this manner.

Eye position dependent spatial tuning

The change in visual responses to retinotopically identical stimuli with eye position could be a result of two possibilities. (1) Cells were coding locations of targets in space independent of eye position. In this situation the cell's receptive field remains static in space and the retinal address of the receptive field would change locations with eye movements to remain constant for spatial location. (2) The receptive fields remain retinotopic and only the responsiveness of the cell varies as a function of eye position. In other words, eye position gates the retinal receptive fields. To distinguish between these two possibilities we mapped entire axes through the center of the receptive fields.

Figure 4 shows examples from four neurons showing angle-of-gaze effects in which the response of the cell is plotted as a function of the retinotopic location of the stimulus along a horizontal or vertical axis through the receptive field for two or more eye positions. It can be seen that the responsiveness varies as a function of eye position, but the peaks and symmetry of the receptive fields do not change. Thus the receptive fields remain retinotopic and it is only the responsiveness of the cells that is modulated by eye position.

We modeled the activity of parietal neurons as a multiplicative interaction of eye position and retinal position using the equation $A = G(e_x, e_y) \times R(r_x, r_y)$, where A is the cells firing rate, G is a gain factor that is a function of horizontal (e_x) and vertical (e_y) eye position, and R is the visual stimulus response profile that is a function of horizontal (r_x) and vertical (r_y) retinal locations. This is a four-dimensional problem which will be simplified to two dimensions of vertical eye position and retinal position to show how this multiplicative interaction can lead to a tuning for locations in space, but dependent on eye position. Figure 5A shows a contour plot of simulated activity for a neuron with a planar gain field, which is a sloping line in one dimension, and a retinal receptive field profile that has been fit with a Gaussian curve along a vertical axis. By making the substitution that location in head-centered coordinates is equal to the sum of retinal position and eye position, the activity is plotted as a function of the location of the visual stimulus in head-centered space along the abscissa and eye position along the ordinate. Figure 5B shows the same plot using actual recording data; the similarity of the two plots indicates that the simple multiplicative model seems to be sufficient to fit the data. Examination of these plots shows that the cell is spatially tuned to give a best response when the visual simulus is located at 20° down in craniotopic space. If the cell's retinal receptive field was not gated by eye position, then the cell would give the same response at every position along a 45° diagonal on the contour

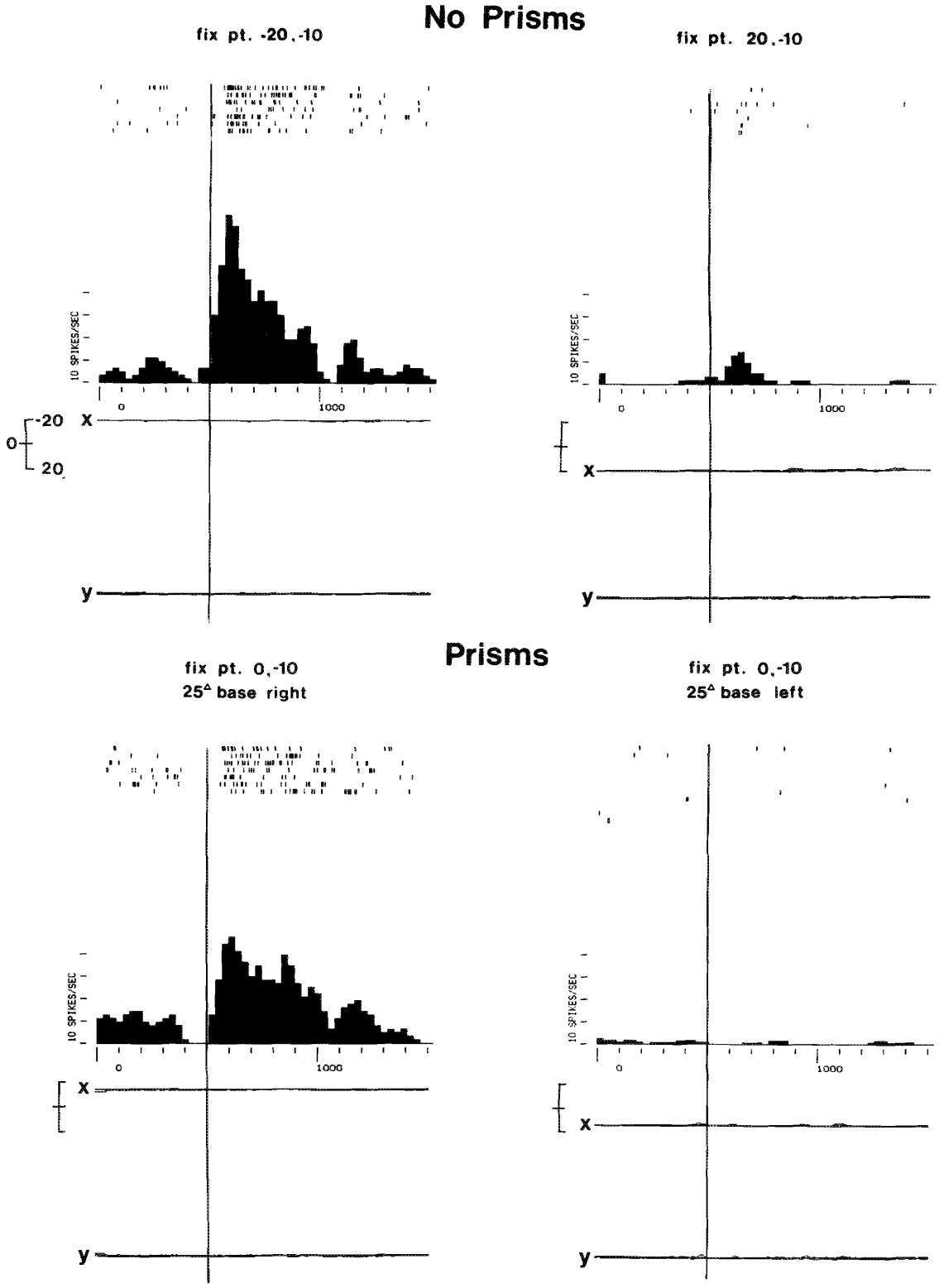

Fig. 3. Control experiment to access the effects of visual background on the eye position tuning of the visual response. These experiments were performed in a lighted room. In the upper left panel a visual stimulus was presented when the animal was fixating 20° left and 10° down from straight ahead. The horizontal (x) and vertical (y) eye position records are shown below the spike rasters and histogram. The upper right panel shows the response for a retinotopically identical visual stimulus but now with the animal fixating to the right 20°. Note that the cell responds better to a retinotopically identical visual stimulus when the animal is looking left compared with right. In the lower panels the animal is made to look left and right by using prisms rather than moving the fixation point on the screen. The eye position recordings indicate that the animal is diverting its gaze by the appropriate amounts to fixate the fixation point. Again the cell responds best for retinotopically identical stimuli when the animal is fixating left as opposed to right. In the upper two panels the retinotopic locations of the visual background of the light room imaged on the retinas is different for looking left and right. However, in the lower two panels, the visual backgrounds are retinotopically identical for fixating left and right since the fixation point is not moved with respect to the background. Since the effects of eye position on the evoked visual response are the same in both the "no prism" and "prism" conditions, the visual background cannot account for this effect and it must be attributed to an eye position signal modulating the visual response.

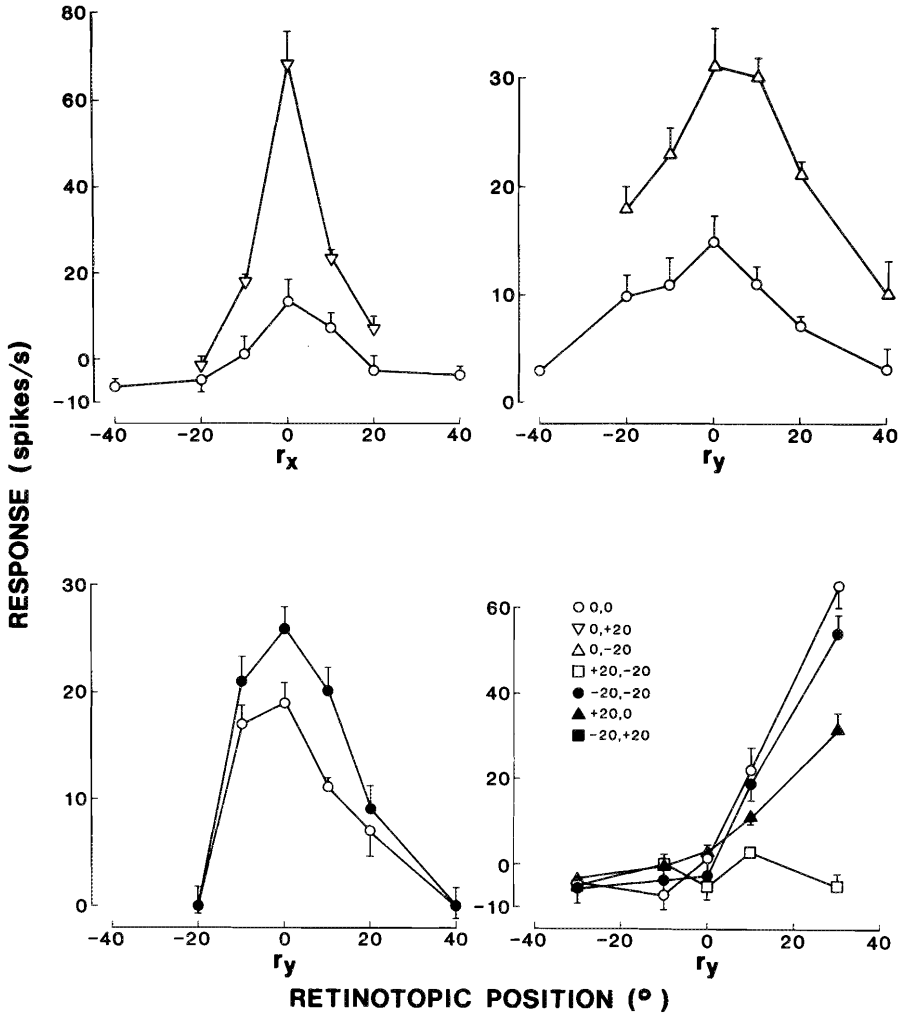

FIG. 4. Mean response rates for different eye positions plotted in retinal coordinates along horizontal (r_x) or vertical (r_y) axes passing through the centers of the receptive fields of four neurons; each graph shows data for one neuron. Each point represents the mean response (±standard error) for eight repetitions of the stimulus presented at the same retinal location. A randomized block design was used to present stimuli to different retinal locations in the receptive field of each cell. The reported response at each retinal location is equal to the activity during the presentation of the stimulus minus the background activity determined before the stimulus presentation. From Andersen et al. (1985), reproduced from Science (Washington, DC), 230: 456–458, © AAAS.

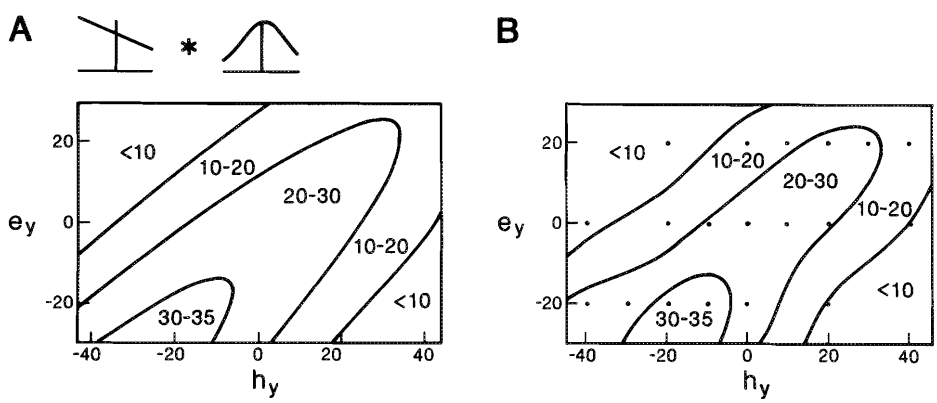

FIG. 5. (A) Computer simulation of the response (in spikes per second) of an area 7a neuron predicted by multiplying the vertical axis of a planar gain field by the vertical axis of a Gaussian receptive field. The results are represented on the contour plot with the stimulus head-centered coordinates (h_y) plotted along the abscissa and eye position (e_y) along the ordinate. (B) Contour plot of actual recording data for a cell with the same gain field and receptive field characteristics as the model neuron plotted in (A). Each data point represents the mean evoked response to eight repetitions of the stimulus; the average standard error for these data points was two spikes per second. From Andersen et al. (1985), reproduced from Science (Washington, DC), 230: 456–458, © AAAS.

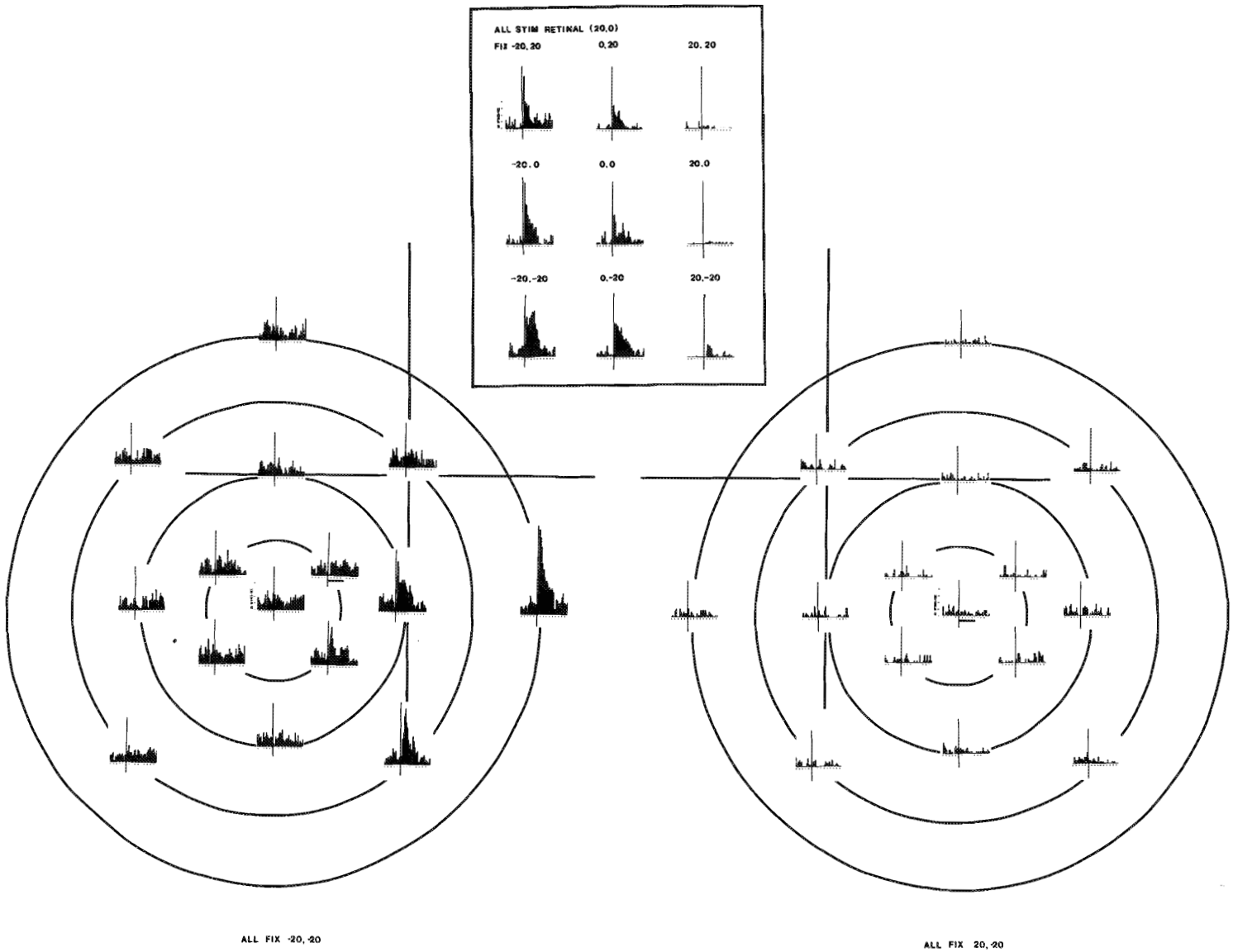

FIG. 6. Plots of activity for stimuli delivered at 17 retinal locations for each of two different eye positions. The box shows the spatial gain field for this cell. The cell is most responsive when the animal looks left and down. The circles represent isoretinal contours at 10, 20, 30, and 40° radii. The large crosses indicate the cardinal axes of the screen coordinates centered on fixation straight ahead. Each histogram is positioned at the retinal locus of stimulation. The data on the left were collected with the animal looking down 20° and to the left 20°; the data on the right were collected with the animal looking down 20° and to the right 20°. For rightward fixation, note that the visual response is not only absent, but that the cell also has a much lower rate of background activity.

plot. However, the plot also shows that the cell only responds to this spatial location when the animal is looking down; in other words, the spatial tuning is eye position dependent. If the tuning were eye position independent, then the cell would give the same response for 20° down in head-centered space at all eye positions.

Problems to be addressed by models of spatial representation in area 7a

We have never found cells that are spatially tuned in an eye position independent manner. Other laboratories have also found visual responses modulated by eye position in other parts of the brain (Aicardi et al. 1987; Funahashi et al. 1985; Peck et al. 1980; Schlag et al. 1980). In those cases where the receptive fields were mapped at different eye positions, which is required to determine the nature of the modulation, a similar eye position dependent tuning was also found. From these experiments it must be concluded that the neural representation of space in an eye position independent manner is distributed. Moreover, the finding that more areas than just the parietal cortex have this type of coding suggests that this type of distributed coding is rather universal for different brain structures. This distributed coding presents a problem: how do you read out the code for spatial location? One way would be to map the locations of spatial tuning systematically across the tangential dimension of area 7a. Recording experiments have not revealed any obvious topography for spatial tuning, indicating that if such a map for craniotopic space does exist, it will at best be crude (Zipser and Andersen 1988). Moreover, the very large receptive fields and spatial tuning fields would tend to mitigate against high resolution mapping like that found for retinotopy in V1.

Another unusual feature of the receptive fields of parietal cells is their complexity (Fig. 11). The fields are large, give approximately equal weighting to the fovea and periphery, and can have multiple peaks. The fields generally have smoothly varying levels for response for nearby sample points. No topographic organization for retinal location of the receptive fields has been found. What is the significance of the shapes of these fields?

Another interesting aspect of the area 7a neurons is that the background activity of the cells also often varies as a function of eye position. Figure 6 shows an example of the most common case in which the background activity varies in the same direction as the gain on the visual response. This cell has a retinal receptive field located to the right and down. The gain field in the middle panel shows that fixations down and to the left produce increased responsiveness to the visual stimulus. On the left and right are more complete maps of the retinal receptive field. The circles represent isoretinal contours in 10° steps and the cross indicates the screen coordinates. For the plot on the left, the retinal receptive field has been mapped with the animal looking down and left; in the panel on the right, the animal is fixating down and to the right. The onset of the visual stimulus is indicated by the vertical line on each histogram and the duration of the stimulus by the bar on one histogram in each panel. Note that when the animal is looking left there are evoked responses for stimuli presented to the right and down, but when the animal is looking to the right there is absolutely no evoked visual response. The major point to note in this figure is that the background activity is high for left fixation when the visual responsiveness is high, and is negligible for right fixations when the visual responsiveness is also negligible.

That the background activity varies in the same manner as the visual sensitivity is not unexpected. For instance, one can imagine that the eye position input increases the visual response by depolarizing the membrane and that in some cases this depolarization not only lowers the cell's threshold but also fires the cell, leading to an increase in its activity. However, some cells showed modulation of the background activity with eye position but no effects on the visual response. And still other cells showed background effects that went in the opposite direction to the gain of the visual response. Again any model of spatial coding in area 7a would need to explain this behavior.

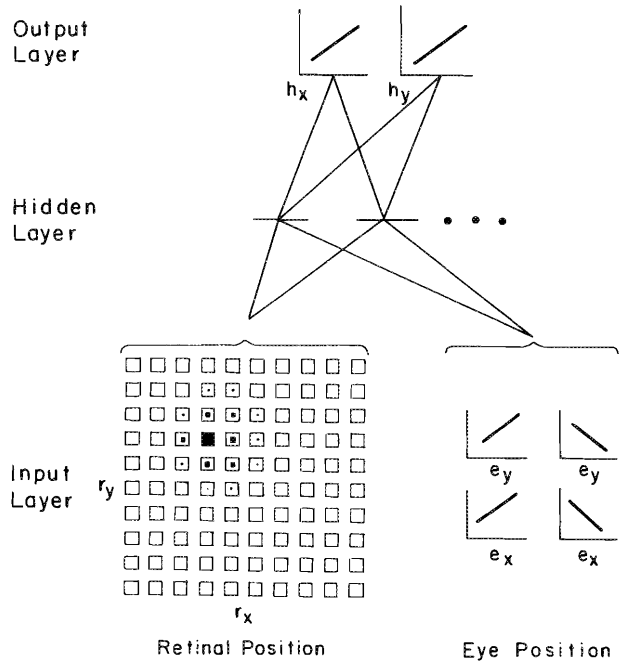

FIG. 7. Schematic of the back-propagation network used to model area 7a. The input to the network consists of retinal position and eye position information. The activity of the output units is a monotonic function of the location of the visual stimulus in craniotopic coordinates. The middle or "hidden" layer units map input to output. The details of the network are explained in the text.

Network model for spatial representation

We have created a parallel network model that learns to map inputs of retinotopic position and eye position to an output of location in head-centered space (Zipser and Andersen 1988). This network consists of three layers and uses the back-propagation learning algorithm (Rumelhart et al. 1986). Interestingly the units in the middle layer that accomplish the spatial transformation show the same eye position dependent spatial tuning properties that are found for posterior parietal neurons. This model also generates retinal receptive fields similar to those found in area 7a neurons and reproduces similar background activities. The remarkable similarity between the model and experiment suggests that the distributed spatial coding discovered in area 7a neurons is indicative of spatial transformations carried out using the same computational algorithm discovered by back-propagation.

Figure 7 shows a schematic diagram of the network. The input layer consists of a 10 × 10 retinal array and four eye position units. The retinal receptive fields are Gaussian in shape with $1/e$ widths of 15°. This input is designed to be similar to the receptive fields found in area 7a that do not show eye-position effects and are assumed to be the retinal inputs in area 7a. The

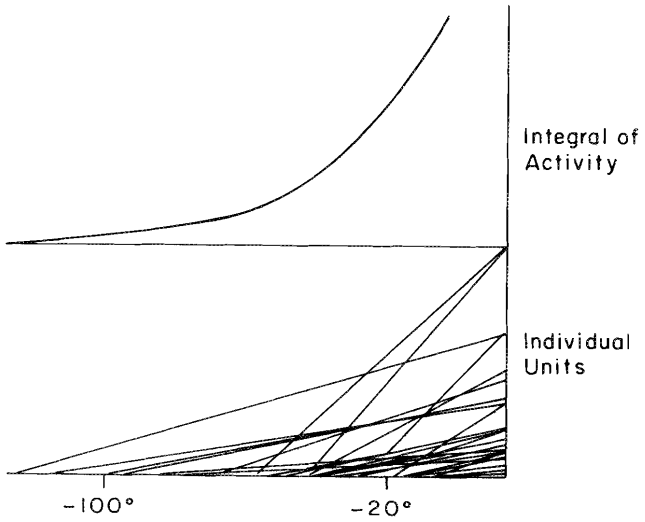

FIG. 8. The lower plot shows linear fits to the horizontal or vertical component of activity of 30 eye position neurons. Plots with negative slopes have been inverted. The upper plot shows the cumulative activity of the 30 neurons.

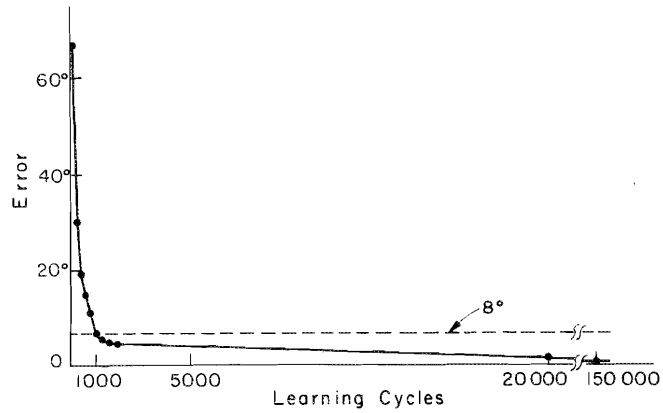

FIG. 9. Plot of the error of the output as a function of the number of training cycles. The distance between the receptive field centers of the retinal input units is 8°; an error of this magnitude is achieved after about 1000 repetitions.

centers of the 100 receptive fields are equally spaced over the 10 × 10 grid with 8° spacings. The four eye position units consist of two units coding a vertical position and two horizontal positions using opposite, symmetrical slopes (Fig. 7). Each unit used either a linear or, in later simulations, a squared function to approximate the signal coming from eye-position cells. The rationale for using a squared function is to approximate the cumulative response of a group of eye position cells; this approach is demonstrated graphically in Fig. 8. Eye position inputs are assumed to be those cells in area 7a that have only eye position signals and no visual response. These cells generally code horizontal and (or) vertical position in a linear fashion (Zipser and Andersen 1988). Measurements were only made to −20°, but the lines are extrapolated back to their intercepts (lower graph of Fig. 8). The upper graph of Fig. 8 is simply the sum of activities of the cells plotted in the lower graph of Fig. 8. As can be seen, a square function would approximate this plot. We ran simulations with both square functions and simple linear functions and got indistinguishable results indicating that the exact representation of the eye position does not appear to be too crucial as long as it is a monotonically increasing function.

The intermediate layer receives inputs from all 104 input units and in turn projects to two or four output units. The output units code position in head-centered space as a linear function of spike rate. There are four outputs units with pairs of opposite slope for horizontal and vertical position. As in the case with the eye-position inputs, the exact form of the function did not appear to matter as long as it was monotonically increasing; also whether only one unit of positive slope or two units of opposite slope were used did not seem to be important. The rationale for using a monotonic function for the output was that the eye position cells could be used as a teaching signal if the animal saccaded to fixate the stimulus. In later experiments we also tried mapping to Gaussian sloped representations of head-centered location with interesting results listed below.

The output of each cell in the network is calculated by first summing all inputs, both inhibitory and excitatory, and then calculating the output as a sigmoidal function of the input. More precisely $o_j = 1/(1 + e^{-(\sum_i w_{ij} o_i + \theta_j)})$, where o_j is the output of a unit in layer j that receives inputs from cells o_i from layer i through synaptic weights, w_{ji}. This sigmoid is chosen as an output function, since it is similar to the operation performed by actual nerve cells that sum inputs, have a threshold, and saturate at high levels of activity. The term, θ_j, is a bias or threshold that can be considered equivalent to the local inhibition within layer j. This bias can be either trained or set and simulations using both of these options will be discussed.

The network begins training with all the connections set at random weights and completes training when the output units accurately indicate positions in head-centered space for any pair of arbitrary retinal and eye position units. The network learns by subtracting the output vector from the desired output vector for each input pattern to generate an error. This error signal is then propagated back through the network to change the weights in the network. The back-propagation algorithm ensures that the weights will change to reduce error in the performance of the network. The actual equations and derivation of the back-propagation procedure are covered in Rumelhart et al. (1986). This cycle is repeated until the network reduces error to desired levels. The spatial transformation network we have constructed learns quickly and always settles to very low error values. Figure 9 shows a training curve for one simulation. Within 1000 trials, the network is already showing accuracy that is better than the spacing of the distance between the centers of the retinal receptive fields. By running the network for a long time, it continues to improve to vanishingly small errors.

After training is complete the middle layer units have receptive fields that remain retinotopic, but their activity becomes modulated by eye position in a manner similar to that seen in the recording data from area 7a neurons. Figure 10 shows receptive fields from four hidden units. Each square in the 3 × 3 array is a map of visual responses to a probe stimulus in retinotopic coordinates. The darkened areas represent firing rates over 50% of the maximum firing rate. Receptive fields are mapped over an area of visual field 80° in diameter; the nine fixation locations are equally spaced within a 40 × 40° box centered on straight ahead. Each square represents the visual field that is mapped and positioned to correspond to the fixation position that was used when it was mapped. Inspection of these graphs indicates that the receptive fields remain retinotopic but that the responsiveness changes with eye position. Moreover, the change in responsiveness is roughly planar and similar to a majority of the gain fields recorded from area 7a neurons. It can

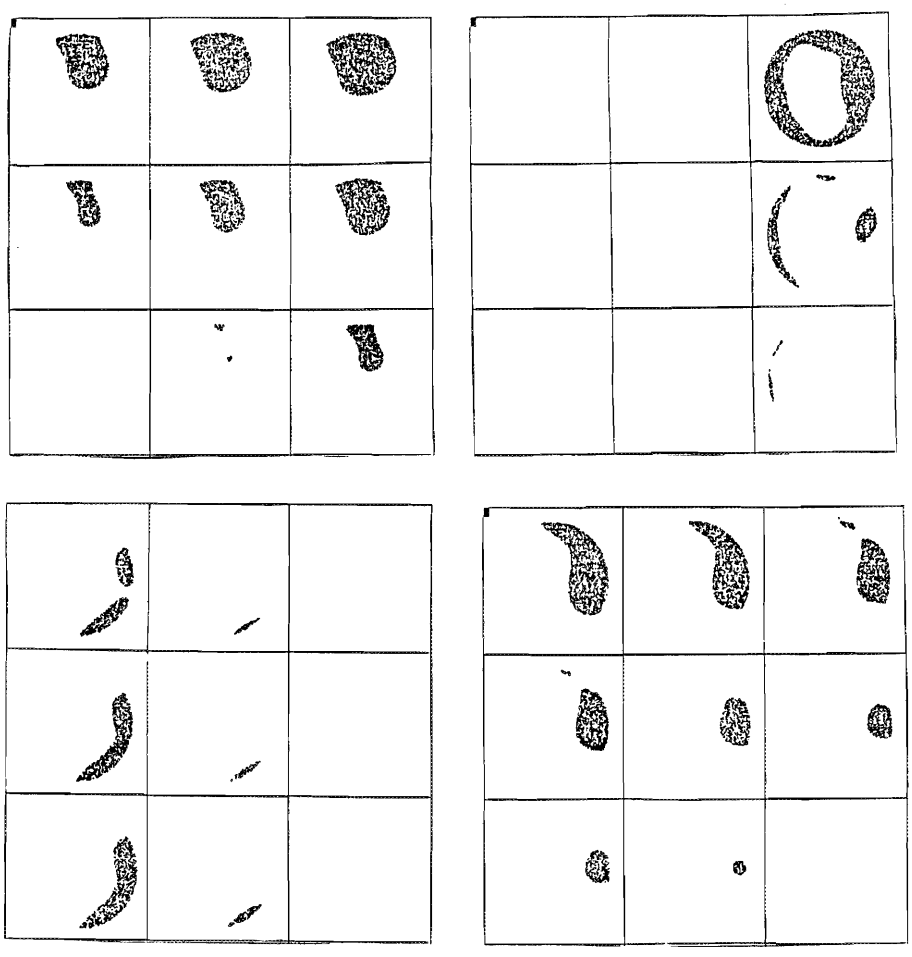

FIG. 10. Plots of receptive field properties for four hidden units after training the network. The nine squares are visual receptive field maps at nine different eye positions. The eye positions are each spaced by 20° of visual angle with the middle square at straight ahead. Each box contains a complete receptive field map over a circular area of the visual field 80° in diameter. The plotting is such that only activity over 0.5 is plotted. Notice that the receptive fields remain retinotopic for the different angles of gaze and only the magnitude of the activity is modulated by eye position. Also note that the gain fields are generally planar, like those seen in the recording data.

also be seen that the receptive fields are large and can have peculiar shapes, not unlike the cells in area 7a.

In Fig. 11 the retinotopic visual receptive fields from recording experiments are compared with retinotopic receptive fields generated by the model. It should be emphasized that these comparisons are intended to be qualitative and only show that they are similar; obviously they will not be exactly the same, just as no two area 7a neuron receptive fields will ever be identical. Surfaces have been fit to the recording and model data using a Gaussian interpolation algorithm. Each receptive field is 80° in diameter. The fields have been categorized by complexity into three classes: class 1 has the simplest receptive fields with each having a single peak of activity; class two fields are of intermediate complexity with each field having a single greatest peak of activity and one or more smaller peaks of activity; class three receptive fields are the most complex with each having multiple peaks of greatest activity. The most complex fields are similar to fields of the untrained model. The trained model seldom produces such complex fields.

Next we compared gain fields generated by the model with data gain fields. Figure 12 shows gain fields for all nine hidden units of a simulation. The nine pairs of circles for each unit represent the activity for the same retinotopic stimulus delivered at nine different eye positions. The dark, inner circle's diameter is proportional to the response to the visual stimulus alone and the outer circle diameter is proportional to the entire activity, both background and evoked activity. In most cases, for instance the upper left unit, the background activity varied in the same direction as the gain on the evoked response. However, in some cases, for instance the lower left unit, the background activity actually varied opposite the evoked activity. Although there are no examples of it in this particular simulation, in some cases the background activity varied with eye position, but the gain of the response remained constant. Thus the same effect of eye position on background activity seen in the recording experiments also appears in the model. A close examination of the logistic function used to compute the output of the neurons explains these effects. When the cells are operating on the low end of the sigmoid where it is accelerating upward, the eye position signal increases the background activity and the combination of the eye position and retinal inputs multiplies at the output to produce the gain which varies with the background. When operating on the linear part of the sigmoid, the eye position signal modulates the background activity; but the retinal and eye position inputs add linearly and thus the gain of the evoked response is constant at all eye positions. When the cell is operating on the upper, saturating part of the sigmoid, large eye position signals drive the cell near saturation and result

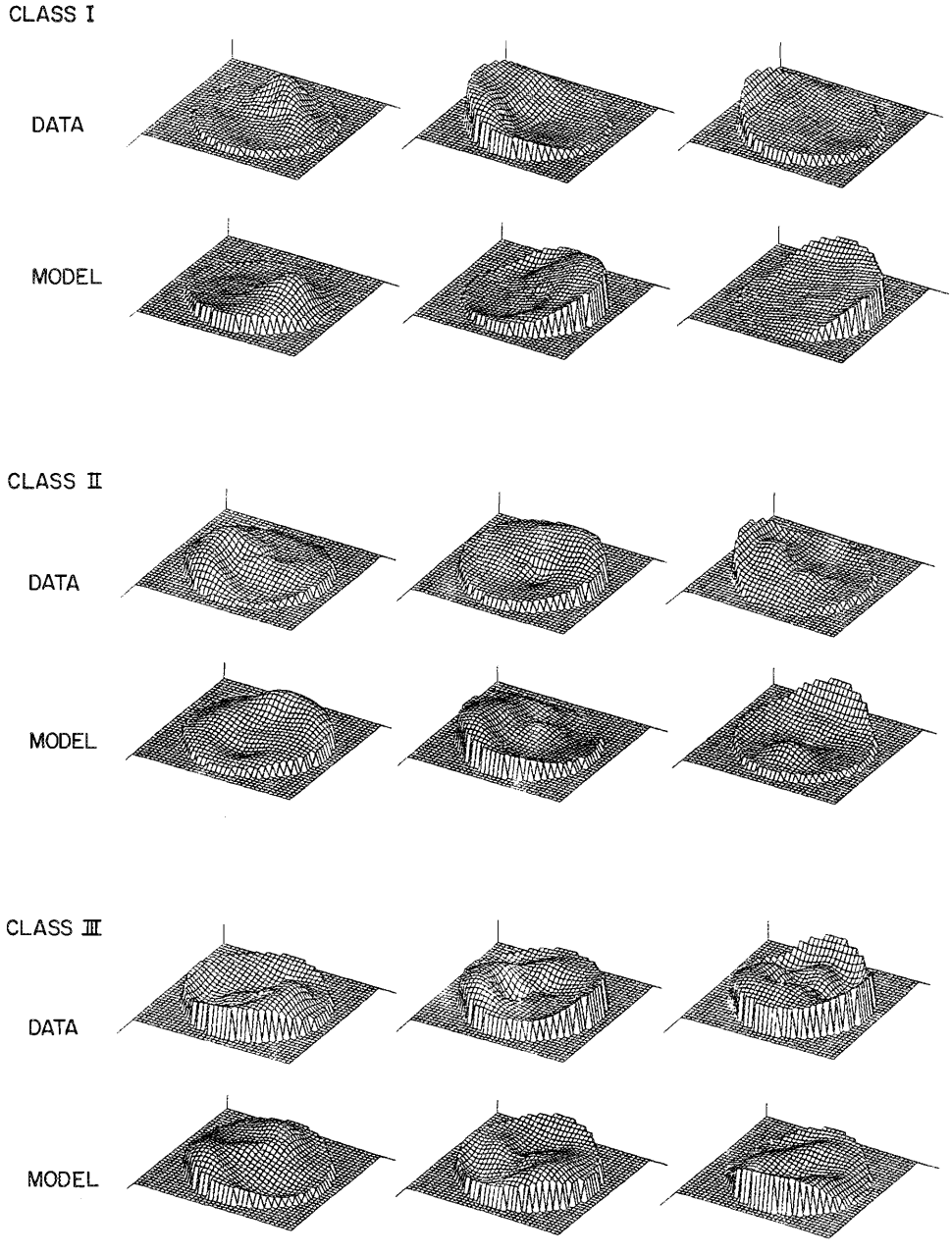

FIG. 11. Visual receptive fields from the data and the model are compared. The receptive fields were divided into three classes: class 1 cells have a single, smooth peak of activity; class 2 cells have one peak of activity but also other smaller peaks or depressions in the receptive field; and class 3 cells have multiple peaks of activity. Note the close correspondence between model and data receptive fields.

in small gains on the visual response; whereas small eye position signals produce larger gains on the visual input because the cell is operating on the nonsaturating portions of the sigmoid. Thus the magnitudes of the visual response change in opposite direction to the background when varying eye position.

In the example given in Fig. 12, the threshold (bias) was free to be trained as well as the synaptic weights. This resulted mainly in the background responses and the evoked responses changing in parallel with eye position. Twenty-eight percent of the cells that showed planar gain fields for the total response (background and evoked response) showed this parallel change. Three examples of this kind of gain field for real neurons are shown in Fig. 13A.

Model simulations in which the threshold (bias) is clamped often produced gain fields in which the visual response changed with eye position, but the background showed little change (Fig. 14). This was particularly true when the background activity was small (Fig. 14). This type of gain field was the most common for the recording data and made up 48% of the planar gain fields. As with the model data they generally had low background activities. Three examples from the recording data are shown in Fig. 13B.

Figure 15 shows the result of running the network when it is mapping to Gaussian receptive fields in head-centered coordinate space rather than mapping to a monotonic function of spatial position. This coding is essentially equivalent to the coding at

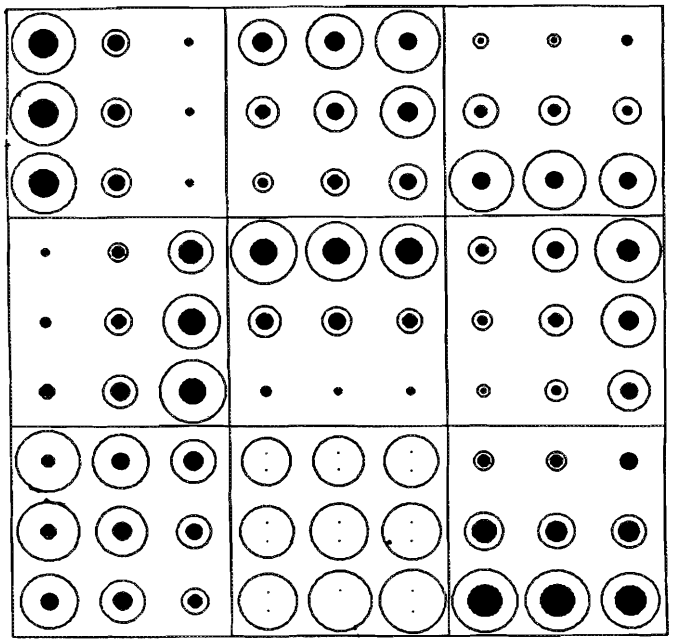

FIG. 12. Gain fields for nine hidden units that were trained to a monotonic output function for location in head-centered space. In this simulation the threshold of each unit is trained as well as the synaptic weights. The diameter of the inner dark circles is proportional to the activity of the visually evoked response, the diameter of the outer circle is proportional to the total activity before subtracting the background, and the size of the annulus is proportional to the amount of background activity. The circles are located in the same relative positions as the fixation positions that produced them. Note that the background activity tends to vary in parallel with the visually evoked response.

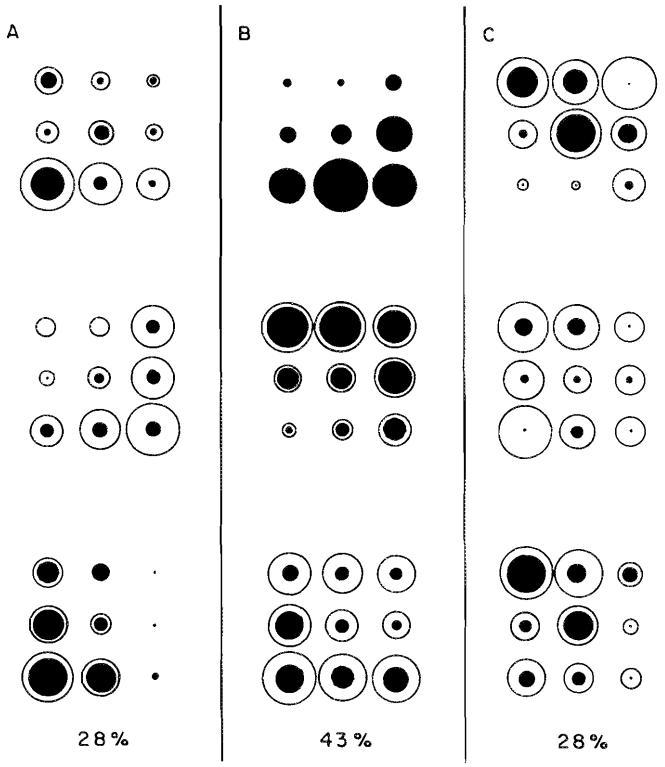

FIG. 13. Examples from the recording data of the three types of planar gain fields defined by the relational variation of background and visually evoked responses with eye position. (A) Background and evoked visual responses change in parallel. (B) Evoked activity changed with eye position but the background activity remained constant. Most of these cells had low rates of background activity. (C) The background and evoked activities changed in different directions but the overall activity remained planar.

the retinal input but with one major difference: the receptive fields are in craniotopic rather than retinotopic coordinates. This mapping results in hidden units whose overall activity is generally planar, but the relative contribution of background and evoked responses varies in a complex and unpredictable manner. Twenty-four percent of the neurons with overall planar gain fields showed this type of variation with three examples illustrated in Fig. 13C.

As can be seen by the above data, the model predicts quite accurately the gain fields, visual receptive fields, and background activity for the majority of cells in area 7a that show planar gain fields. The model did not, however, produce the peaked gain fields that were found in 22% of the cells from the recording data. These cells may not be involved in the coordinate transformation that we have modelled. In the mapping to a monotonic output, hidden units similar to the cells in 13c were found, but much less frequently than when the mapping was to Gaussian, craniotopic receptive fields. This raises the possibility that two types of output representation are mapped by the posterior parietal cortex. Finally, comparing the effects of clamping or training the threshold (bias) levels suggest that about half the cells receive lateral inhibitory inputs that are not adjusted by learning.

Discussion of the model

The simulation results show that training a parallel network to perform coordinate transformations produces the same type of distributed code found in area 7a. This similarity is so interesting that it should be pursued to determine if it represents

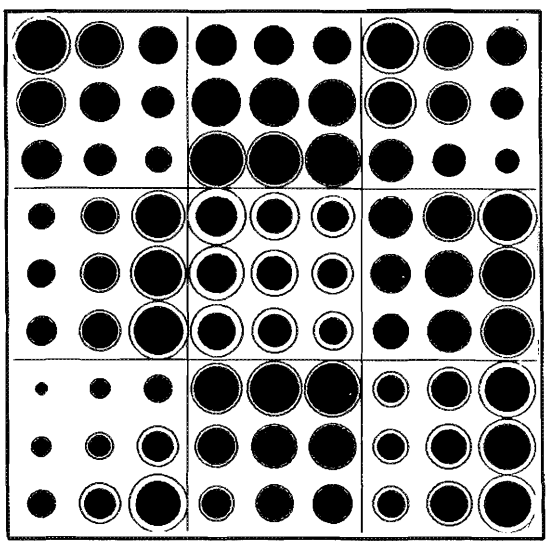

FIG. 14. Gain fields for nine hidden units trained to a monotonic output function for location in craniotopic coordinates, but with the thresholds fixed. Note that the threshold is set such that there is very little background activity and minimal variation of that activity with eye position.

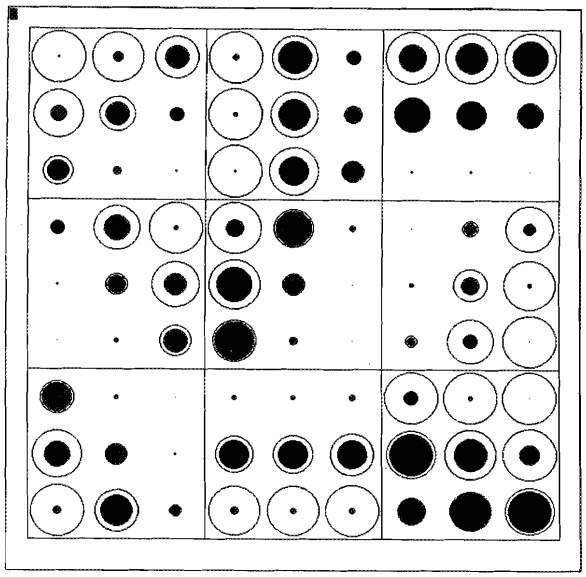

FIG. 15. Gain fields for nine hidden units trained to a Gaussian output function for location in head-centered coordinates. Note that the background activities and evoked activities vary in a seemingly unrelated fashion, but the sum of their activities still produces planar gain fields.

a fundamental outcome of using parallel networks to perform coordinate transformations. One possible line of research would be to make the model more complex by incorporating more features analogous to those found in the brain such as Hebb learning and reciprocal pathways for error feedback. If these more complex models still produce the same distributed code, then this type of coding is very interesting indeed. Another avenue would be to see if this model generalizes to three-dimensional space and body-centered coordinates by collecting data under these conditions for parietal neurons and comparing the results with predictions from the model.

These results suggest that the posterior parietal cortex learns to associate body position with respect to visual space. Thus the parietal lobe appears to form associative memories to generate a look-up table for performing spatial transformations. A learning theory for parietal spatial functions seems to make sense, since it would not be practical to hard-wire spatial representations during development when the body is changing size. Moreover, distortions of space with prisms lead to rapid recalibration in adults, suggesting that unlike ocular dominance there is no critical period for spatial representations and they remain plastic in adults.

It is important to note that the model by definition does not have a topographic organization. Thus there is no requirement for topographic organization in the brain. The reason topography is not necessary is that the organization of the network is distributed and the information is contained in the weights in the synapses. It would be interesting to determine whether putting a crude topography into the connections of the network would accelerate learning.

If the spatial representation in area 7a lacks topography, this does not of course mean that there is not topographic organization in this area. We imagine that learning spatial localization can occur within the dimensions of a typical cortical column; i.e., 1 mm^2. Recordings made within an area of parietal cortex of this size contain a complete complement of receptive fields, eye position signals, and gain fields necessary for a complete representation of craniotopic space (R. A. Andersen, G. K. Essick, and R. M. Siegel, unpublished observation). Thus spatial location can be mapped over and over again in many repeating units in the cortex and may overlay some as yet unknown functional repetitive architecture that would need in each of its modules the complete machinery for coordinate transformations.

The rather large receptive fields suggest that every posterior parietal neuron has access to the entire retina and the particular shape of the receptive fields is due to competition during learning. It is interesting that this competition produces in both the parietal neurons and the model units receptive fields that, although they are complex, tend to coalesce so that they are smoothly varying rather than random in structure. The large receptive fields are presumably due to cascading divergence that occurs at each stage in the cortico-cortical projection from V1 to area 7a.

The fact that many fewer eye position synapses than retinal synapses were required at the convergence onto the hidden units in the model has interesting parallels to the anatomy of the parietal lobe connections. It is believed that the source of eye position information comes from lower brainstem centers and is relayed through the intralaminar nuclei (Schlag-Rey and Schlag 1984). However, these nuclei are small compared with the cortical areas relaying visual information to area 7a.

There is the question of where the output units of the model might exist in the brain. As mentioned earlier, cells showing the expected eye position independent behavior have not been found in area 7a. One possibility is that these cells exist in another brain area that receives input from area 7a. Another possibility is that this distibuted coding is carried and used throughout all brain regions that need spatial representations. The final spatial output might be seen only in the output of the physical plant. For instance, the cells innervating the eye muscles are broadly tuned and the position of the eye is course coded in a distributed fashion over the activity of the six extraocular muscles. Thus the final spatial output may be pointing the eye or a finger accurately to a location in space and no single cell in the brain might be found that codes the location of visual space in an eye position independent fashion.

Finally there is the quesiton of whether this form of model of distributed coding in parallel networks, which appears to explain the parietal data rather well, will be useful in other brain regions. Recently Maunsell and Schiller examined the response of area V4 cells in a task in which the monkey must match the orientation of a visual or somatosensory cue grating with the orientation of a visual test grating (J. H. R. Maunsell, personal communication). They found that cells that respond to the cue and cells that respond to the test stimulus are orientation tuned, and that cells show facilitated activity for a particular combination of cue and test stimulus. Thus the activity of some V4 neurons shows a multiplicative interaction between two inputs that is similar to the interaction for eye and retinal inputs for area 7a neurons. It would be interesting to construct a similar network in which the inputs were the cue and test stimuli and the output the correct match. Would the hidden units develop properties like those of V4 neurons? Similar interactions between cue and stimulus have also been noted in the prefrontal cortex in delayed response tasks (Watanabe 1987).

Acknowledgements

We thank Carol Andersen for editorial assistance. R.A.A. was supported by grants EY05522 and EY07492 from the

National Institutes of Health, the Sloan Foundation, and the Whitaker Health Sciences Foundation. D.Z. was supported by grants from the System Development Foundation and the Air Force Office of Scientific Research.

AICARDI, G., BATTAGLINI, P. P., and GALLETTI, G. 1987. The angle of gaze influences the responses to visual stimulation of cells in the V3-complex of macaque monkey visual cortex. J. Physiol. (London), **390**: 271.

ANDERSEN, R. A. 1987. Inferior parietal lobule function in spatial perception and visuomotor integration. *In* Handbook of physiology. *Edited by* F. Plum and V. B. Mountcastle. American Physiological Society, Rockville, MD. pp. 483–518.

ANDERSEN, R. A., ESSICK, G. K., and SIEGEL, R. M. 1985. Encoding of spatial location by posterior parietal neurons. Science (Washington, DC). **230**: 456–458.

FUNAHASHI, S., BRUCE, C. J., and GOLDMAN-RAKIC, P. S. 1985. Visual properties of prefrontal cortical neurons. Soc. Neurosci. Abstr. **11**: 525.

HALLET, P. E., and LIGHTSTONE, A. D. 1976. Saccadic eye movements toward stimuli triggered by prior saccades. Vision Res. **16**: 99.

MAYS, L. E., and SPARKS, D. L. 1980. Dissociation of visual and saccade-related responses in superior colliculus neurons. J. Neurophysiol. **43**: 207–232.

PECK, C. K., SCHLAG-REY, M., and SCHLAG, J. 1980. Visuo-oculomotor properties of cells in the superior colliculus of the alert cat. J. Comp. Neurol. **194**: 97–116.

ROBINSON, D. A. 1975. Oculomotor control signals. *In* Basic mechanisms of oculor motility and their clinical implications. *Edited by* P. Bach-y-Rita and G. Lernerstrand. Pergamon, London. pp. 337–374.

RUMELHART, D. E., HINTON, G. E., and WILLIAMS, R. J. 1986. Learning internal representations by error propagation. *In* Parallel distributed processing: explorations in the microstructure of cognition. Vol. 1. Foundations. *Edited by* D. E. Rumelhart, J. L. McClelland, and the PDP Research Group. Bradford Books/MIT Press, Cambridge, MA. pp. 318–362.

SCHLAG, J., SCHLAG-REY, M., PECK, C. K., and JOSEPH, J. P. 1980. Visual responses of thalamic neurons depending on the direction of gaze and the position of targets in space. Exp. Brain Res. **40**: 170–184.

SCHLAG-REY, M., and SCHLAG, J. 1984. Visuomotor functions of central thalamus in monkey. I. Unit activity related to spontaneous eye movements. J. Neurophysiol. **51**: 1149–1174.

WATANABE, M. 1987. Activity of primate prefrontal neurons related to complex behavioral task. Soc. Neurosci. Abstr. **22**: S579.

ZIPSER, D., and ANDERSEN, R. A. 1988. Back propagation learning simulates response properties of a subset of posterior parietal neurons. Nature (London), **331**: 679–684.