A Neural Network Model of Flexible Spatial Updating

Robert L. White III and Lawrence H. Snyder

Department of Anatomy and Neurobiology, Washington University School of Medicine, St. Louis, Missouri 63110

White, Robert L. III and Lawrence H. Snyder. A neural network model of flexible spatial updating. *J Neurophysiol* 91: 1608–1619, 2004. First published December 10, 2003; 10.1152/jn.00277.2003. Neurons in many cortical areas involved in visuospatial processing represent remembered spatial information in retinotopic coordinates. During a gaze shift, the retinotopic representation of a target location that is fixed in the world (world-fixed reference frame) must be updated, whereas the representation of a target fixed relative to the center of gaze (gaze-fixed) must remain constant. To investigate how such computations might be performed, we trained a 3-layer recurrent neural network to store and update a spatial location based on a gaze perturbation signal, and to do so flexibly based on a contextual cue. The network produced an accurate readout of target position when cued to either reference frame, but was less precise when updating was performed. This output mimics the pattern of behavior seen in animals performing a similar task. We tested whether updating would preferentially use gaze position or gaze velocity signals, and found that the network strongly preferred velocity for updating world-fixed targets. Furthermore, we found that gaze position gain fields were not present when velocity signals were available for updating. These results have implications for how updating is performed in the brain.

INTRODUCTION

When our gaze shifts, images of objects move on the retina in predictable ways. The image of an object that is fixed in the world will shift on the retina in proportion to the magnitude and direction of a gaze shift. However, not all objects are fixed in the world. Many objects move and, in particular, many objects move with the gaze of the observer. For example, when a moving object is tracked, all features of the tracked object, both foveated and nonfoveated, maintain their locations on the retina. We can say that a nonmoving object is stationary in a world-fixed reference frame, whereas a moving target that is being tracked by gaze is stationary in a gaze-fixed reference frame. If a target becomes occluded, then the ability to maintain the location of the target in memory requires that its frame of reference be taken into account.

A number of cortical areas store remembered spatial locations (Funahashi et al. 1989, 1990; Gnadt and Andersen 1988). A simple means of storing world-fixed locations would be an explicit world-centered representation. However, many visuospatial representations are gaze-centered (e.g., Colby et al. 1995; but see Duhamel et al. 1997; Galletti et al. 1993; Olson 2003), and behavioral evidence supports the use of gaze-centered representations to store world-fixed locations (Baker et al. 2003). For a gaze-centered representation to encode information about a world-fixed location, the representation must be updated each time gaze shifts. There is evidence for such updating in many cortical and subcortical areas (Batista

and Andersen 2001; Bruce and Goldberg 1990; Duhamel et al. 1992; Mazzoni et al. 1996; Nakumura and Colby 2000, 2002; Walker et al. 1995). For example, neurons in the lateral intraparietal (LIP) area in the posterior parietal cortex encode the goal location of an impending saccade (Gnadt and Andersen 1988). If an intervening saccade to a new fixation point is performed between target presentation and the final saccade to the target, neurons in LIP update their activity to represent the oculocentric coordinates of the final goal relative to the new eye position (Duhamel et al. 1992; Mazzoni et al. 1996). As a result of this updating, neurons in LIP encode information about targets that are world-fixed without explicit world-centered encoding. Updating occurs not only in response to saccades, but to any gaze perturbation, including whole body rotation [with vestibuloocular reflex (VOR) suppression] and smooth pursuit eye movements (Baker et al. 2002; Powell and Goldberg 1997).

What mechanisms underlie the capacity of gaze-centered cortical areas to compensate for shifts in gaze? The receptive fields of LIP neurons are fixed with respect to the retina (Colby et al. 1995), but their visual responsiveness (gain) can be modulated by changes in gaze position. This modulation is called a gaze position gain field (Andersen and Mountcastle 1983). By virtue of combining eye position information with retinotopic information, gain fields may provide an implicit head-centered representation of visuospatial information (Zipser and Andersen 1988). These gain fields could be used in connection with updating in a double-step saccade task (Xing and Andersen 2000a).

Although updating is appropriate for world-fixed targets, it is inappropriate for gaze-fixed targets. Separate neural networks could be responsible for encoding spatial memories in world-fixed and gaze-fixed frames of reference. Alternatively, spatial memories could be computed by a single network that appropriately updates memories based on a particular reference frame.

To explore the computational basis of updating, we created a neural network model using a 3-layer recurrent architecture and trained it to flexibly update contingent on whether a target was world-fixed or gaze-fixed. Neural network models have been used to approximate the input–output relationships of neural circuits in the brain (reviewed in Zipser 1992). The internal behavior of such models often mimics the properties of single neurons in the brain (Mitchell and Zipser 2001; Xing and Andersen 2000a,b; Zipser and Andersen 1988). We asked whether a simple neural network model was capable of performing spatial updating in a reference frame-dependent manner. We then compared the output of the model with the

---

behavior of macaques performing a similar task (Baker et al. 2003). Finally, we probed the internal representation of the model to make predictions about how neurons in the brain might implement the computations under consideration, focusing on the roles of gaze position signals, gaze velocity signals, and gaze position gain fields.

We found that a simple recurrent network could successfully compensate for or ignore gaze perturbation signals based on a contextual cue. When we compared the model to monkeys performing a similar task, we found that both animal and model were less precise in generating world-fixed output. Units in the hidden layer of the network performed a distributed computation: individual units were not specialized for performing only the gaze-fixed or world-fixed task, but in fact could discriminate well on both types of trials. The model could perform the updating task with either a gaze position or gaze velocity signal, but performed better with and would self-organize to select the gaze velocity signal. Finally, we found that the presence of gain fields in the model's hidden layer was dependent on the type of gaze signal used by the network. Gain fields were present only in networks that relied on position signals to perform the updating task.

METHODS

We designed a neural network model containing units in 3 layers (input, hidden, and output). Figure 1 depicts a schematic of the model's structure. The input layer is designed to represent a visual map in one dimension of space and to provide information about gaze position and gaze velocity. Output is in eye-centered coordinates and can be read out as encoding the goal location of an upcoming saccade. For flexible behavior, we added a cue that indicated whether the target was world-fixed or gaze-fixed. The hidden layer is given recurrent connections to provide the network with memory over time. Units in each layer can take on activity values from 0 to 1 (except where stated otherwise). The network has full feedforward connections: every unit in the input layer projects to every unit in the hidden layer, and every unit in the hidden layer projects to every unit in the output layer. Each connection has an associated weight that indicates the strength of that connection. The activity of an individual unit in the hidden or output layer is determined by the logistic function $f(x) = 1/(1 - e^{-x})$, where x is the weighted sum of the inputs plus a bias. In addition to the feedforward connections, the network contains recurrent connections between every unit in the hidden layer. Thus the input to a hidden unit at a discrete time t includes the feedforward connections from the input layer at time t plus the recurrent connections from the hidden
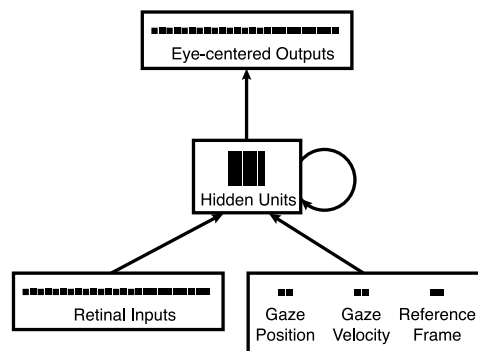

FIG. 1. Architecture of the recurrent network model. Network has 3 layers (input, hidden, and output). Input layer includes retinal inputs (one-dimensional) and extraretinal inputs (gaze position, gaze velocity, and reference frame cue). Hidden layer has fully recurrent connections. Output layer encodes saccadic goal location in eye-centered coordinates.

layer at time $t - 1$. For an individual trial simulation (see *Task and training*), the activity of the network is calculated for 13 consecutive time steps. We define each step to be 100 ms, for ease of comparison with physiological data. Individual weights were initially set to random values between -0.1 and 0.1.

In the input layer, the retinal array is modeled as 25 visual units with receptive fields evenly distributed over a range, defined as -60 to $60°$. Each retinal unit has a Gaussian receptive field with a $1/e^2$ width of $7°$. Gaze position is linearly encoded by 2 units. Activity of the first position unit is scaled such that 0 activity corresponds to straight ahead $(0°)$ and the activity range $[-1, +1]$ corresponds to gaze position in the range $[-40°, 40°]$. The second position unit has the opposite activity of the first (a simple push–pull model). A second pair of push–pull units linearly encodes velocity. Velocity units can encode a range of gaze speeds between -200 and $200°/s$. Finally, 2 binary reference frame units with opposite polarity indicate whether the output should be world- or gaze-fixed. The representation of the cue signal in the cortex is likely to be complex, involving not only sensory input but also prior experience and current expectations. The model presented here does not address these complexities, but instead considers only how flexible visuospatial processing is accomplished. To that end, we chose a very simple representation of the cue: a binary switch. The hidden layer consists of between 5 and 65 units (see RESULTS). Most analyses were performed with 25 units in the hidden layer. The output layer consists of an array of 25 units encoding saccadic goal location in oculocentric coordinates. The units are evenly distributed over the range -60 to $60°$ and have Gaussian receptive fields identical to the visual input units.

Task and training

The task of the network was to store and if necessary transform a pattern of activity representing the spatial location of a target. To correctly perform the task, the network had to either compensate for or ignore any changes in gaze that occurred during the storage interval, depending on the instruction provided by the 2 reference frame input units. The network output provides an eye-centered representation of the stored spatial location, which can be interpreted as a location in a salience map (Colby and Goldberg 1999) or as a goal location for an upcoming saccade (Snyder et al. 2000). For this particular network, these 2 descriptions are interchangeable.

A target was presented for one time step, or 100 ms, at the onset of each trial. A target could appear at 1 of 9 locations, evenly spaced within the central one third $(40°)$ of the workspace. The reference frame cue, which determined whether gaze perturbations should be incorporated into the output or ignored, was present for the entire duration of the trial. Gaze position at the start of the trial was at one of 4 locations $(-15, -5, 5, \text{or } 15°)$. Memory-period gaze perturbations were 10 or $20°/s$ to the left or right, were initiated 300 ms after the target disappeared, and lasted for 500 ms. The network output (saccade amplitude or target location) was read out from the network 0 to 400 ms after the end of the gaze perturbation.

In an analogous task, 2 rhesus monkeys performed memory-guided saccades to world-fixed and gaze-fixed target locations after whole body rotations (VOR cancellation), smooth pursuit eye movements, or saccades (see Baker et al. 2003). Slow gaze perturbations (rotations, pursuit) were $10°/s$ to the left or right from fixation points at -20, -10, 0, 10, or $20°$ and lasted for either 300 or 600 ms. Monkeys were cued $400-1,200$ ms after the end of the gaze perturbation to make a saccade to the remembered target location.

In the modeling experiments, the source of the gaze perturbation signal (e.g., rotation, pursuit, saccade) is relevant only inasmuch as different sources might have different properties. Three properties are of particular interest: *1*) whether the source provides gaze position or gaze velocity information, *2*) whether the gaze perturbation extends over a long or short period of time, and *3*) whether the endpoint of the change is known at the start of the movement (e.g., saccades). Here we focus on gaze position

and gaze velocity signals and briefly address duration of movement. We discuss the third property in the discussion.

The network was often grossly inaccurate when the saccadic goal lay on the boundaries of the workspace (defined by the output). To minimize these edge effects, trial types requiring saccades $>20°$ in either reference frame were excluded. This left 96 remaining combinations of target location, eye position, and gaze perturbation vectors that constituted the training set.

The network was trained using the "backpropagation through time" algorithm (Rumelhart et al. 1986; Williams and Zipser 1995). The algorithm optimizes the connection weights between units to minimize the error produced at the output. An additional constraint was enforced during training: the hidden-to-output weights were constrained to values greater than -0.1. Forcing positive weights to the output makes the hidden units more likely to develop response fields similar to those of the output units (Mitchell and Zipser 2001). Removing this constraint did not change our overall conclusions, however.

From our experience training monkeys on an identical task (Baker et al. 2003), we suspected that training the network in stages might facilitate learning, rather than presenting the complete task and training set from the very start. This suspicion was confirmed by preliminary studies. Networks presented with the entire training set from the start of training converged slowly or converged to local minima (unpublished observations). Therefore we chose a graduated training regime analogous to that used for the monkeys.

The network was initially trained using a simple saccade task: no gaze perturbation and no memory period. Once this was mastered, a short (100-ms) memory period was introduced, and then gradually lengthened to 1,200 ms. Weights were adjusted after every cycle of the complete training set until an error threshold was reached. The threshold was set at a high value to avoid overtraining the network on the memory saccade task. Again, this is analogous to our experience in training monkeys; we do not dwell too long on any one stage, else achieving the next stage becomes more difficult (unpublished observations). With the network, each time the duration of the memory period was increased, smaller and fewer weight updates needed to be performed to reach the threshold. Hence, the learning rate started at $\eta = 0.05$ and was decreased inversely with the number of steps.

Once it could perform memory saccades, the network was introduced to the full task. Training had an equal probability of occurring on any one of the 4 postperturbation time steps or on the time step directly preceding the onset of gaze perturbation. The latter case promotes stability in the network output and enforces the condition that world-fixed and gaze-fixed trials should produce the same output if no perturbation has occurred. Training was not performed *during* the gaze perturbation to avoid enforcing a particular time course of output during the perturbation period (a post hoc comparison revealed slightly improved performance when training also occurred during the perturbation, but no other differences). Weights were adjusted after each complete cycle through the training set. Training proceeded for 5,000 cycles at a learning rate of $\eta = 0.001$. Subsequent training for 7,500 cycles occurred with the learning rate halved every 2,500 cycles. This process of simulated annealing helps avoid both local minima (by using an initially high learning rate) and limit cycles (by moving to a low learning rate) and generally helps optimize both the speed and the accuracy of training (Mitchell 1996).

ROC analysis

A receiver operating characteristic (ROC) was calculated for each unit in the hidden layer in the manner of Britten et al. (1992). The receiver discriminated between 2 conditions [whether a target was inside or outside of the response field (RF)] based on the activity of the unit. For a given criterion level of activity, the proportion of trials on which the inside-RF response exceeded the criterion was plotted against the proportion of trials on which the outside-RF response exceeded the criterion. Points were calculated for 3,334 criterion

values uniformly distributed over the range of possible activity values ([0, 1]). The collection of points forms a stair-step function on the unit square from (0, 0) to (1, 1). The area under this curve (AUC) was calculated. The AUC describes how well an ideal observer could correctly discriminate between the inside-RF and outside-RF conditions based on the responses of the unit.

RESULTS

Performance of the model

We constructed and trained a recurrent 3-layer neural network to flexibly perform a spatial updating task based on a contextual cue. The neural network (Fig. 1) received information from simulated retinal and extraretinal inputs. The retinal inputs encoded the position of a visual stimulus in one dimension of space. The extraretinal inputs provided signals proportional to gaze position and gaze velocity, as well as a binary reference frame cue (either world-fixed or gaze-fixed). The task required the network to store and later report the location of the target. Like the representations of many visual and oculomotor cortical areas (Colby et al. 1995), both the network input and output are eye-centered. As a result, the representation of a world-fixed target must be modified to compensate for an intervening gaze perturbation, whereas the representation of a gaze-fixed target remains unchanged (Fig. 2).

The trained network was able to successfully report the location of both gaze-fixed and world-fixed target locations. Figure 3A shows the inputs and outputs for an example world-fixed trial. The target appears 15° to the right of the center of gaze and then disappears. A rightward 20° gaze perturbation displaces the location of the target, measured with respect to the center of gaze, from 15° right to 5° left. The output units of the network reflect this displacement over time (Fig. 3A, *lower panel*). In a matched trial using the same initial target location and gaze shift but a gaze-fixed rather than world-fixed target, the target remains at 15° to the right and the output of the network remains constant over time (Fig. 3B, *lower panel*).

To quantify the network's performance, saccade location was read out as the center of mass of the output units (Fig. 3C). Each unit "votes" for its location in the array, weighted by its activity. The sum of these weighted outputs was compared with the correct world-fixed or gaze-fixed target location (Fig. 3C). For each combination of target location and perturbation, we compared the network's ability to distinguish gaze-fixed versus world-fixed target locations by constructing a modulation index (MI):

$$\text{MI} = (\text{saccade amplitude} - \text{gaze fixed target location})/(\text{gaze perturbation})$$

If the network performed perfectly, $\text{MI} = 1$ for each world-fixed target and $\text{MI} = 0$ for each gaze-fixed target. The mean MI for world-fixed targets was 0.97 compared with 0.06 for gaze-fixed targets. The distribution of indices for the entire training set is shown in Fig. 3D. It is clear that the network effectively discriminates between gaze-fixed and world-fixed conditions and correctly localizes the memorized target in the appropriate reference frame.

We computed the root-mean-squared (RMS) error for the all combinations of target locations and gaze perturbations for world-fixed and gaze-fixed trials. RMS error varied with the number of hidden units, but error on world-fixed trials was always greater than that on gaze-fixed trials (Fig. 4). Error as a function of the number of hidden units described a U-shaped function, with the
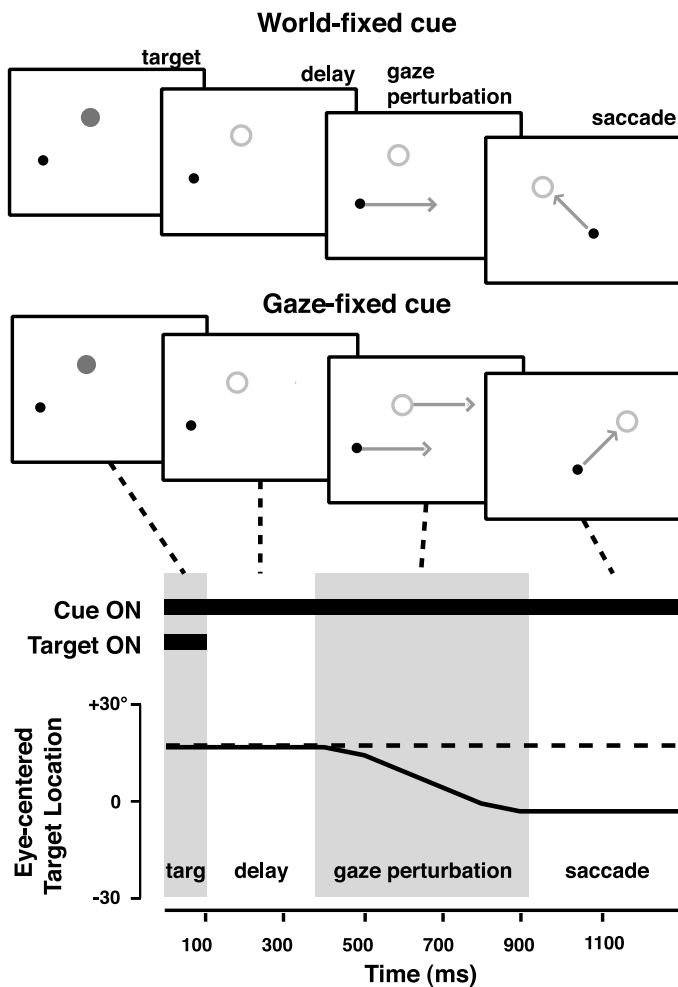
World-fixed cue

Gaze-fixed cue

FIG. 2.   Schematic of the task conditions. Relative locations of the target and fixation point are illustrated in a 2-dimensional workspace, although the model tracked only the horizontal component. Network was given a transient visual target ("Target ON") and a sustained reference frame cue ("Cue ON"). A slow gaze perturbation followed, after which the network was queried for a saccade vector to either the world-fixed (solid line) or gaze-fixed (dashed line) location of the original target.

least error occurring between 20 and 40 units. Networks with fewer than 5 hidden units did not converge on a solution within 12,500 cycles, and networks with >40 hidden units also performed increasingly poorly. The performance of networks with >40 hidden units was improved by training over more cycles or with larger learning rates, whereas the performance of networks with fewer than 40 hidden units was not significantly altered with these manipulations. For a network with 25 hidden units, the mean RMS error across 3 simulations was 1.93° for world-fixed targets and 1.19° for gaze-fixed targets. The data shown in Figs. 3 and 5–7 are taken from a network with 25 hidden units, but varying this number within the range of 20–40 did not affect our conclusions.

Network output correlates with animal behavior

Network inputs and outputs were chosen to be biologically plausible representations in an effort to model what may occur in monkey posterior parietal cortex (Duhamel et al. 1992; Mazzoni et al. 1996; Xing and Andersen 2000a). To justify comparisons between model units and neurons, we first com-

pared network output with animal behavior on a similar task. Monkeys accurately directed saccades to the remembered locations of world- and gaze-fixed targets after the direction of gaze was shifted during a memory period (Baker et al. 2003). Gaze was perturbed by either a whole body rotation, smooth pursuit eye movement, or visually guided saccade. The precision of memory-guided saccades after the gaze perturbation differed depending on whether slow (rotation, pursuit) or visually guided saccadic perturbations were performed. Memory-guided saccades were less precise for world-fixed compared with gaze-fixed targets after slow gaze perturbations, but equally precise after saccadic gaze perturbations.

To compare network and monkey performance, we measured the accuracy of the network by comparing the coded location during the postperturbation interval to the correct target location for all gaze perturbations and starting positions (Fig. 5A, *left panel*). The network is accurate in directing saccades to the appropriate target but, like the animals (Fig. 5A, *right panel*), was hypometric for the most peripheral (20°) target locations. Note that decreased accuracy and decreased variability for the most peripheral target locations (±20°) are not at the edges of the input/output range (±60°).

To extract a measure of variability from the network, it is necessary to introduce noise. We introduced Gaussian white noise (SD $\sigma_n = 0.05$) into both the position and velocity inputs of a trained network and measured the variability of the resulting output (Fig. 5B, *left panel*) over 15 repetitions. The output variability ranged from 2 to 6° and, as with the monkey performing the task with slow gaze perturbations, was greater for world-fixed compared with gaze-fixed saccades. The variability of world-fixed saccades was smaller for the most peripheral targets in monkeys and the network. Thus the network reproduces the patterns of both accuracy and variability seen in the monkey when a whole body rotation or a smooth pursuit eye movement perturbs gaze, but not when gaze is perturbed by a saccade.

When the monkey's gaze was perturbed by a saccade, memory precision was equal for world-fixed and gaze-fixed trials. The difference in the patterns of behavior seen after saccades compared with smooth pursuit or rotations might be attributable to a difference in the duration the gaze change. However, when gaze was shifted in a single time step (100 ms), the pattern of model output remained the same (data not shown).

The difference in the patterns of behavior with saccades compared with pursuit or rotation might also be attributable to a difference in the form of the gaze signal. We therefore tested whether a gaze displacement signal, rather than a gaze position or velocity signal, could approximate the behavior seen in the animal with saccades. A displacement signal differs from a velocity signal in that it encodes a position error (the difference between the initial and goal positions) and not a rate of change; it differs from a position signal in that it encodes only relative position (referenced to the initial position) that resets at the end of each gaze perturbation, such that the goal position then becomes the new initial position when the perturbation is complete. When the network was trained with the displacement signal, variability of memory-guided saccades to world-fixed targets was greater than the variability of saccades to gaze-fixed targets (data not shown), exactly as in the simulated pursuit and rotation tasks. The result was the same regardless of whether the entire displacement occurred in a single time
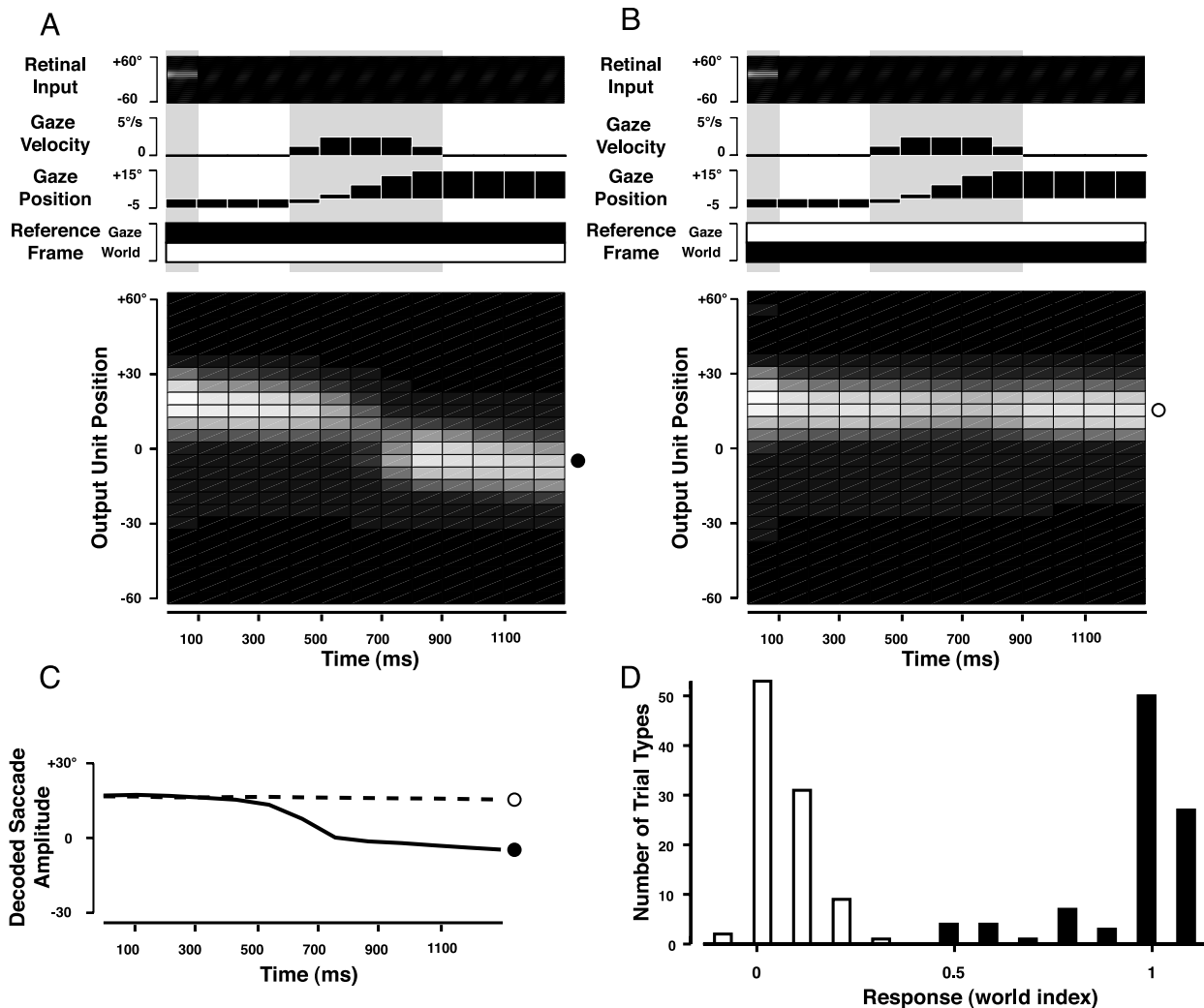
FIG. 3. *A*: network input and output for a single world-fixed trial. *Top*: activation of the retinal input layer, velocity and position signals, and reference frame activations at each time step. *Bottom*: activation of the output layer at each time step. *B*: network input and output for a single gaze-fixed trial (same target and gaze shift as in *A*). In both *A* and *B*, coded location of the correct output (● or ○) is shown for comparison. *C*: saccade locations encoded by the output layer for the trials shown in *A* (solid line) and *B* (dashed line). Coded locations of the correct world-fixed (●) and gaze-fixed (○) outputs are shown. *D*: histogram of modulation index scores (see text) for all the trial types in the training set (white bars, gaze-fixed trial types; black bars, gaze-fixed trial types).

step or across 5 time steps. Thus training with the displacement signal or the position/velocity signal both failed to reproduce the pattern of animal behavior seen with saccades. We conclude that the network reproduces the pattern of behavior after a slow gaze perturbation, and shortening the time course of the gaze shift or providing a displacement signal does not affect the network output.

Hidden units contribute to both gaze-fixed and world-fixed transformations

Although the network is capable of tracking both world-fixed and gaze-fixed targets, the possibility exists that training creates 2 distinct populations of hidden layers units, each of which contributes exclusively to either the gaze-fixed or the world-fixed transformation. Alternatively, a single uniform population of hidden units could contribute equally to both transformations. To assess whether the network's processing was segregated or uniform, we examined the response properties of single units in the hidden layer.

Many hidden layer units updated their activity in response to the gaze perturbation on world-fixed trials, but not on gaze-fixed trials. An example of such a unit is shown in Fig. 6. When a target is flashed in the unit's response field (RF; the region of the retinal input layer that maximally activates the unit) in the context of the gaze-fixed cue, the unit maintains its activity throughout the trial (Fig. 6*A*). If the target is flashed outside the RF, the unit does not respond (Fig. 6*B*). On world-fixed trials in which the target is initially flashed inside the RF, the unit's activity decreases as the gaze perturbation moves the remembered target location outside the RF (Fig. 6*C*). Inversely, when the target is flashed outside the RF the unit increases its activity as the gaze perturbation brings the remembered target location into the RF (Fig. 6*D*).

To quantify the relative amounts of gaze-fixed and world-fixed information each unit conveys, we measured the ROC of an ideal observer of a hidden unit's activity. We introduced a fixed amount of noise to all of the inputs ($\sigma_n = 0.25$) and measured the activity of each unit for 15 repetitions of selected
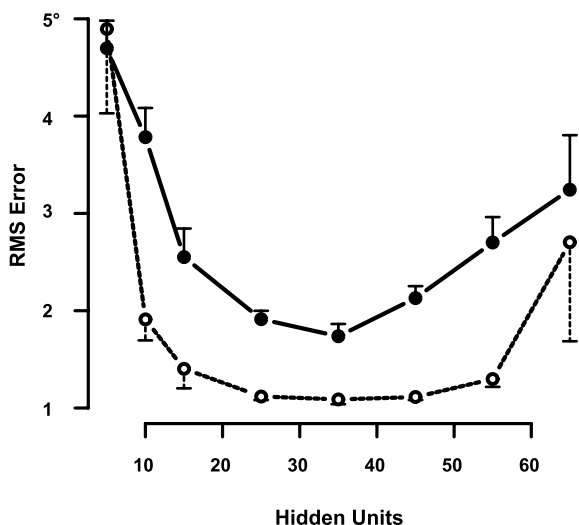
FIG. 4. Performance of networks as a function of the number of hidden units. RMS error on all world-fixed (●) and gaze-fixed (○) trials in the training set for identically trained networks varied with the number of units in the recurrent hidden layer. Best performance was seen with networks with 25 to 45 hidden units. Vertical bars represent SD for 3 networks.

trial types. Trial types were selected to be those in which the target was flashed either inside or outside the RF, followed by a 20° gaze perturbation that brought the target outside or inside the RF, respectively. We constructed our receiver to discriminate whether a target was either inside or outside of the RF (see METHODS). We examined the receiver's performance by calculating the area under the ROC curve (AUC). Perfect discrimination corresponds to an AUC of 1. Chance performance corresponds to an AUC of 0.5. A comparison of the AUC for gaze-fixed versus world-fixed trials reveals that individual hidden units carry different amounts of information in the 2 tasks (Fig. 7).

If the hidden layer had segregated during training into 2 populations, we would expect to see 2 clusters of units in Fig. 7: one cluster in the lower right quadrant that discriminates well only on gaze-fixed trials (gaze-fixed AUC near 1, but world-fixed AUC below 0.5) and another cluster in the upper left quadrant that discriminates well only on world-fixed trials

(world-fixed AUC near 1, but gaze-fixed AUC below 0.5). Instead, the population is unimodally distributed and performs better than chance for both the gaze-fixed and world-fixed tasks.

Network updates using gaze position, gaze velocity, or gaze displacement signals, but strongly prefers velocity

An important question is whether gaze position or gaze velocity signals are used to update spatial memories. This is a difficult question to address in the animal because one cannot easily decorrelate the 2 signals, nor selectively eliminate one input while leaving the other intact. However, with the model, we can train the network with position, velocity, or displacement inputs and compare their performance. We can also train a network with 2 inputs (e.g., both position and velocity) and then selectively lesion either set of inputs after training is complete. From the performance of the lesioned model we can determine whether the network self-organizes during training to rely in part or in whole on just one input or the other. We will refer to this as the network "preferring" or "choosing" one input over the other. We tested the preference of networks for position, velocity, or displacement inputs. To assess the relative preference for one input over another, we presented a maximum of 2 signals during training.

With only position information available during training (velocity inputs set to zero) the network converged on an acceptable solution. The network also converged when only velocity information or displacement information was available (position inputs set to zero). RMS error under world-fixed conditions was statistically identical for displacement and velocity networks (Student's t-test; $P > 0.05$, $n = 3$). In contrast, RMS error was 2-fold greater when the network was forced to rely on gaze position signals (Fig. 8A). Thus although the network can be trained to use any input, performance is superior with velocity or displacement information when compared with position information.

The removal of either input after training with both position and velocity also indicated a strong preference for velocity
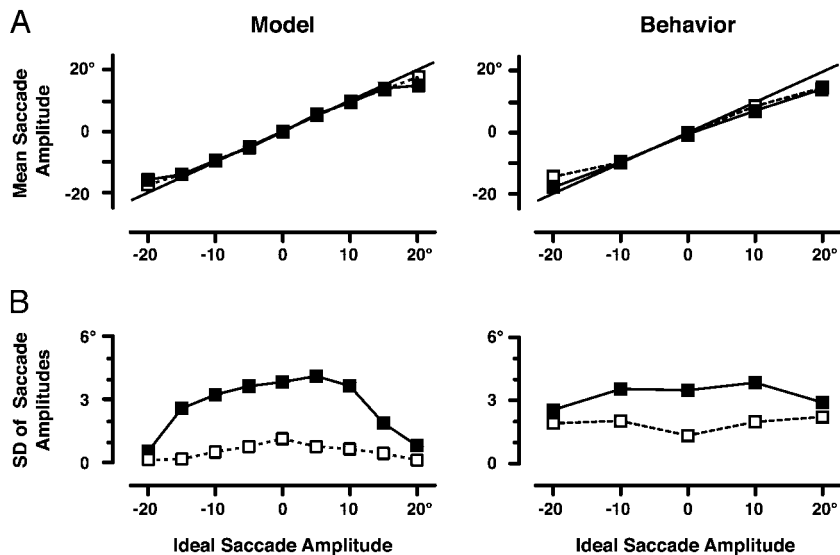

FIG. 5. A: accuracy of network output (*left*) and animal behavior (*right*). *Left panel*: coded output of the network at the last time step vs. the ideal saccade amplitude for world-fixed (■, solid line) and gaze-fixed (□, dashed line) trials. *Right panel*: mean horizontal saccade amplitudes vs. the ideal saccade amplitudes from 2 monkeys (Baker et al. 2002). Straight line indicates perfect performance (actual = ideal). B: precision of network output and animal saccades, measured as the SD of the saccade amplitudes, for the same data as shown in A.
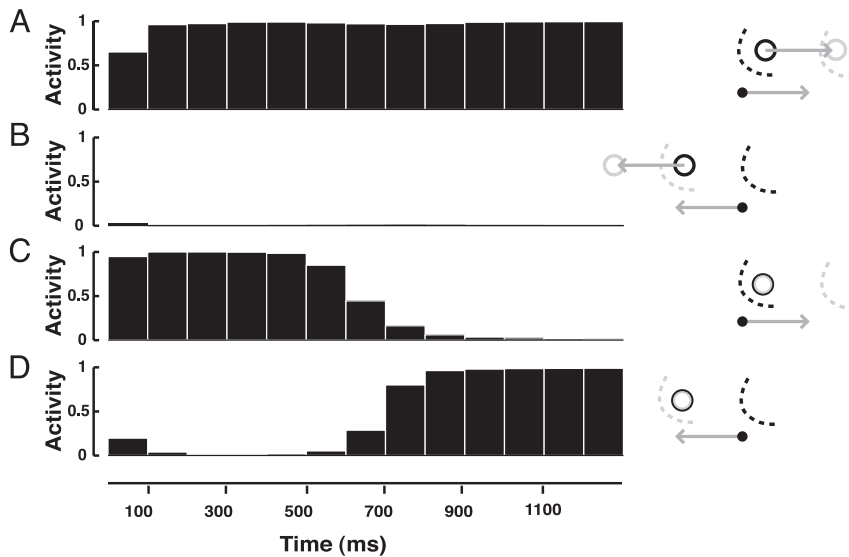
FIG. 6. *A*: activity of one hidden layer unit over time in a single trial when presented with the gaze-fixed cue and target inside its response field (RF). *B*: activity of the same unit when the target is flashed outside the RF in the context of the gaze-fixed cue. *C*: activity of the unit in a world-fixed trial in which the target initially appears inside the unit's RF. Subsequent gaze perturbation moves the remembered target location outside the RF. *D*: activity of the unit in a world-fixed trial in which the target appears outside the unit's RF, but the gaze perturbation brings the remembered target location into the RF.

over position. Although the performance of the network suffered little when the position signal was removed, the world-fixed error increased 5-fold when velocity information was removed (Fig. 8*B*). Thus when provided with both inputs during training, the network preferentially relies on velocity. Note that removing either signal either before (Fig. 8*A*) or after training (Fig. 8*B*) had little effect on gaze-fixed performance. This is consistent with the hypothesis that the network ignored gaze shift information when keeping spatial memories fixed with respect to gaze.

When trained with both displacement and velocity, the performance of the network decreased markedly when either the displacement or the velocity input was removed (Fig. 8*C*). Thus the network relied specifically on the combined input and did not prefer either displacement or velocity in this case. Similarly, networks trained with both displacement and position failed without the combined input (Fig.
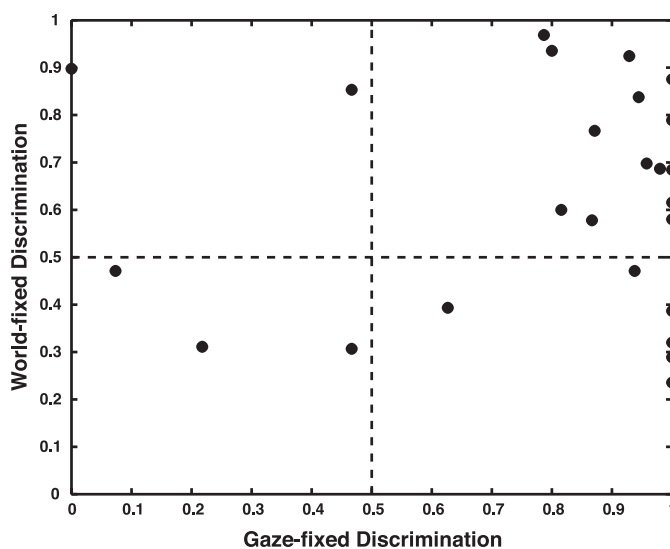

FIG. 7. Comparison of gaze-fixed vs. world-fixed discrimination for units in the hidden layer (*n* = 25). For each unit, the area under the receiver operating characteristic (ROC) curve (AUC) was calculated (see METHODS). Each unit is plotted with respect to its AUC for gaze-fixed (ordinate) and world-fixed (abscissa) trials. Perfect discrimination corresponds to an AUC of 1. Dashed lines indicate the level of chance performance (AUC of 0.5).

8*D*), showing in this case no preference between position and displacement signals.

What causes the clear preference for velocity over position inputs? If the network prefers velocity because of an absolute constraint of the network architecture, then we would expect that it would consistently make the same choice regardless of the quality of the velocity information. Alternatively, if the preference reflects a relative advantage in using velocity compared with position, then changing the relative reliability of the 2 signals should shift the preference of the network.

We manipulated the reliability of gaze information by introducing variable amounts of noise to the position and velocity inputs. To select a suitable amount of noise, we added equal amounts of noise to both of these inputs and measured network performance (Fig. 8*E*, *inset*). It can be seen that the noise degrades performance on the world-fixed task in a log-linear fashion when σ_n is >0.01. Increasing noise on gaze-fixed trials has no effect on performance. A value of total noise $\sigma_n = 0.1$ (arrow) was selected for manipulation of the relative noise to ensure that such manipulation would have observable effects on network performance.

To assess how noise during training might affect the network's choice of gaze signal, we modulated the ratio of noise between velocity and position inputs [noise ratio (NR) = $\sigma_{\mathrm{vel}}/\sigma_{\mathrm{pos}}$]. The sum of the noise was fixed at 0.1 and the noise ratio was varied between 0.01 ($\sigma_{\mathrm{pos}} = 0.01$; $\sigma_{\mathrm{vel}} = 0.09$) and 9 ($\sigma_{\mathrm{pos}} = 0.09$; $\sigma_{\mathrm{vel}} = 0.01$). Figure 8*E* shows the network's performance as a function of the noise ratio. The intact network (triangles) performed better when the noise ratio favored a cleaner velocity signal. When position information was removed after training (hollow squares), the network was only modestly affected when trained with position noise greater than velocity noise (NR <1), but performance worsened as velocity noise increased. When velocity was removed after training (solid squares), performance was severely affected, but more so at low NR values. These data imply that the network's preference to use velocity rather than position information reflects a relative advantage of using velocity information and not an absolute constraint.

Gaze position gain fields are not present in a networks that rely on velocity or displacement

A number of visual and oculomotor-related areas in monkey posterior parietal cortex have been implicated in the updating of spatial information in response to gaze perturbations (Duhamel et al. 1992; Mazzoni et al. 1996; Nakamura and Colby 2000, 2002). One computational theory regarding how updating is accomplished relies on the occurrence of gaze position

gain fields, that is, modulation of responses by eye position. A previous study showed that hidden units activities were modulated by gaze position in a network trained to update target position in a double-step saccade task (Xing and Andersen 2000a). We used our network model to explore whether gaze position gain fields arise whenever retinotopic signals are updated to account for gaze perturbations, or if gaze position gain fields are specific to networks that rely on position information. We tested for gaze position gain fields in the hidden layer of networks trained using gaze position, gaze velocity, or both as input. After training, individual hidden units were tested for modulation of visual responses by gaze position. Gain fields were observed in our network when gaze position information was available. An example of a unit with a gain field is shown in Fig. 9*A*. The unit has a spatially tuned visual receptive field at 0° that remains roughly constant at gaze positions from -15 to $+15°$ (*left panel*). However, the magnitude of the peak visual response varies linearly with gaze position (*right panel*).

Gain field strength was measured as the change in peak response per change in gaze position by linear regression analysis. The distributions of gain field strengths in networks trained with various gaze signals (Fig. 9*B*) revealed that velocity and displacement networks had similar distributions at or near zero, whereas position networks contained units with a broad range of gain field strengths. Networks trained with both position and velocity signals showed an intermediate distribution.

We measured the number of units with gain fields in networks trained with displacement, velocity, position, or combined position and velocity gaze signals. A criterion value of 0.2%/degree (Fig. 9*B*, solid line) was chosen based on published data such as Brotchie et al. (1995) and Snyder et al. (1998); using different criterion values yielded similar results. When only gaze position information was available, 35% of units showed gain fields stronger than 0.2%/degree. When both position and velocity were available, 12% of units had gain fields stronger that 0.2%/degree. However, when networks were trained with velocity alone, only 4% of units displayed gaze position gain fields, despite the fact that these networks performed updating as well as or better than networks trained with only position information. In networks trained with the
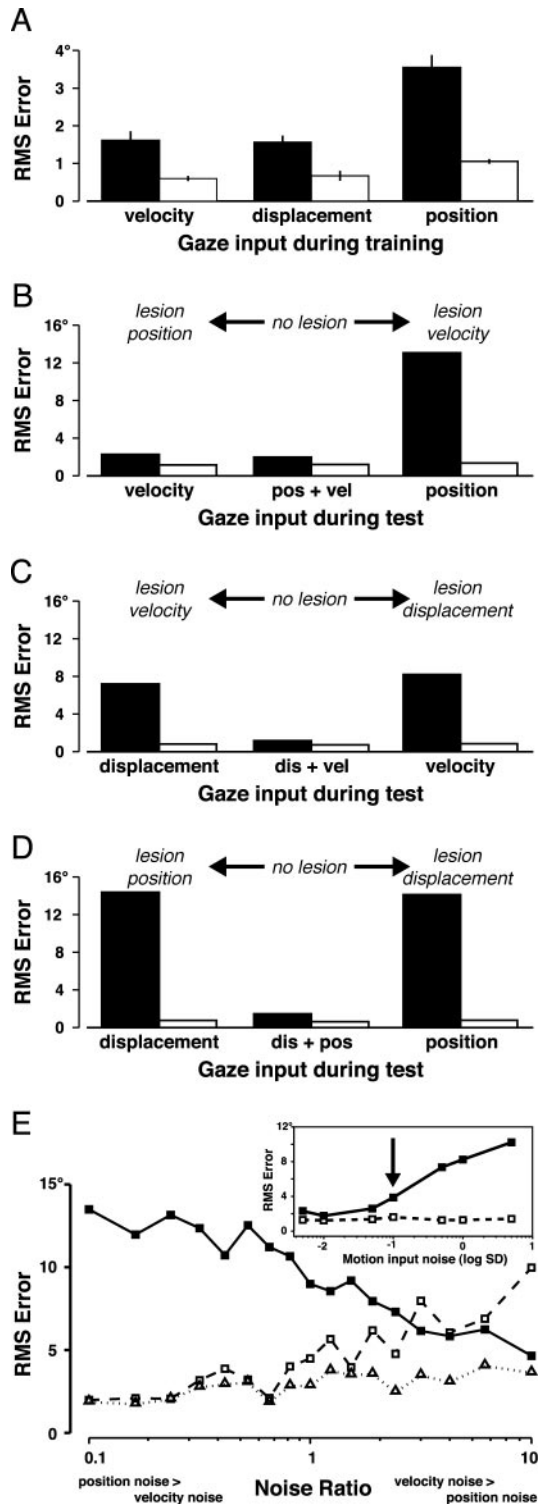

FIG. 8. *A*: performance of the network when trained with only velocity, displacement, or position gaze inputs. Root-mean-squared (RMS) error for all world-fixed (black) and gaze-fixed (white) trial types is shown for networks trained with only velocity inputs, displacement inputs, or position inputs. *B*: performance of the position + velocity network when gaze signals are removed. RMS error is shown for the network with position inputs removed, the intact network, and the network with velocity inputs removed. Network has a strong preference for the velocity input over the position input. Note the change in ordinate scale from *A*. *C*: performance of the displacement + velocity network when gaze signals are removed. Network does not show a strong preference. When either signal is removed, the network fails under world-fixed conditions. *D*: performance of the displacement + position network when gaze signals are removed. Network does not show a strong preference. When either signal is removed, the network fails under world-fixed conditions. *E*: performance of networks trained with varying velocity:position noise ratios. RMS error for world-fixed trials is shown for intact networks (△, dotted line), the same networks with position inputs removed (□, dashed line), and the same networks with velocity inputs removed (■, solid line). *Inset*: performance as a function of gaze noise. RMS error for the same network in *B* vs. the SD of Gaussian noise added to both position and velocity inputs (solid line, world-fixed trials; dashed line, gaze-fixed trials). Arrow indicates the total noise value selected for networks in *E*.
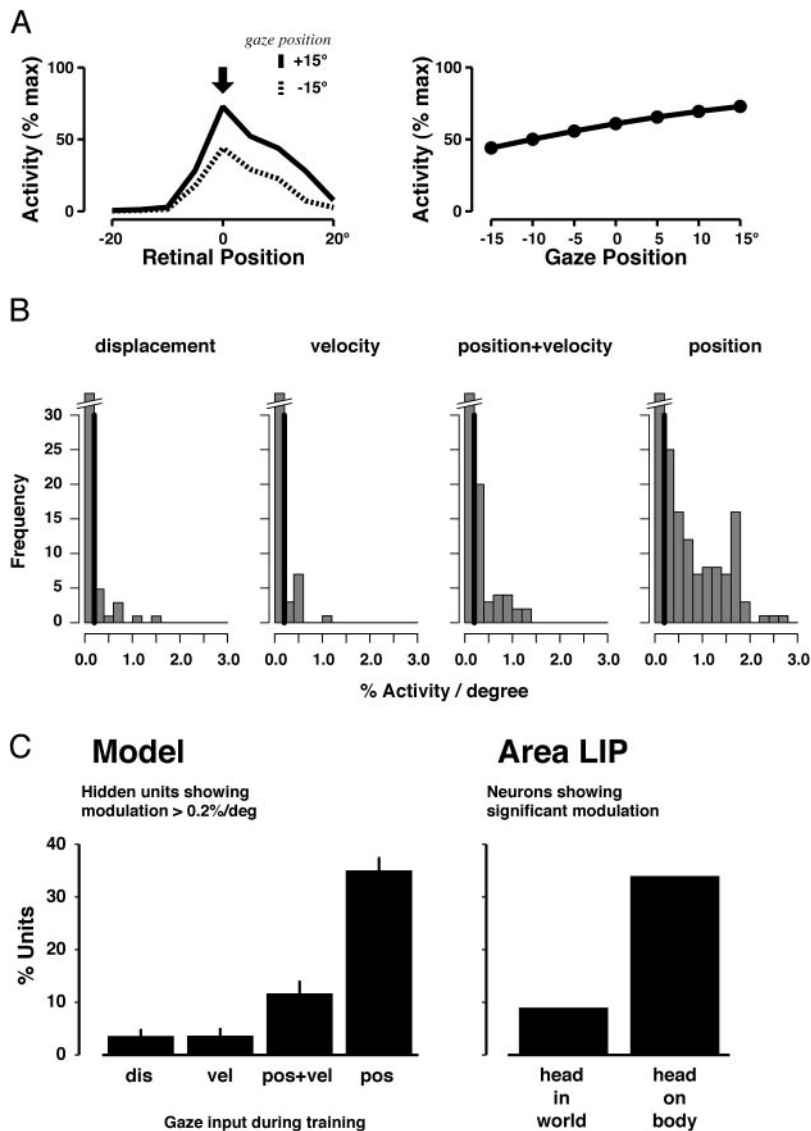
FIG. 9. *A*: example gain field from a single unit in the hidden layer. *Left panel*: unit's response to targets flashed at various retinal positions reveals a spatially tuned receptive field with a peak response at 0° (arrow). This response is modulated by the absolute gaze position (−15° and +15° gaze positions shown). *Right panel*: modulation of peak response is approximately linear across a range of gaze positions. *B*: histogram of gain field strengths measured in networks trained with 4 different gaze input conditions. Black line indicates the criterion value used in *C*. Height of the first bin has been truncated in each graph; numbers of observations in the first bin are (from *left to right*) 289, 289, 265, and 195. For each input condition, a total of 300 observations were made from 6 networks, 25 hidden units per network, and 2 conditions (gaze-fixed, world-fixed) per unit. *C*, *left panel*: proportion of hidden layer units demonstrating gaze position gain fields stronger than 0.2%/degree under the different gaze input conditions. *Right panel*: proportion of single lateral intraparietal (LIP) neurons demonstrating significant body-referenced or world-referenced gain field modulation (Snyder et al. 1998).

displacement signal, only 4% of hidden units showed a gain field modulation.

These results suggest that gain fields are associated specifically with the use of gaze position signals to update spatial information, rather than being a general feature of all networks that perform updating. We hypothesize that a gain field representation is present in networks that rely on position, but not in those that rely on velocity or displacement. The small number of gain fields in networks trained using both velocity and position (12%) suggests a small contribution from the position inputs.

We compared gain fields in the network to those of single neurons in LIP (Fig. 9*C*; neuronal data from Snyder et al. 1998). Gain field modulations were measured for both head-on-body position (orientation of the head on the body) and head-in-world position (orientation of the head relative to an external reference). Many LIP neurons (34%, *n* = 59) showed significant modulations for gaze position referenced to the body (Student's *t*-test, *P* < 0.05), whereas only 10 of 109 cells (9%) were significantly modulated by gaze position referenced to the world. Despite the scarcity of gaze-in-the-world gain

fields, spatial representations of world-fixed targets are nonetheless updated in LIP after passive whole body rotations when gaze rotates with the body (VOR suppression paradigm; Baker et al. 2002; Powell and Goldberg 1997). These data imply that gaze position information is not used for this task. Instead it seems probable that, like our network model, LIP may use gaze velocity signals to update world-fixed target locations in response to whole-body rotation.

DISCUSSION

The neural network model described in this report combines retinal and extraretinal signals to encode the locations of targets that are either fixed in the world or fixed with respect to gaze. The network is able to flexibly switch between the 2 frames, based on a contextual input. The output of the network can be interpreted either as the goal location for an upcoming saccade (Andersen et al. 1992) or as the current locus of attention in eye-centered coordinates (Colby et al. 1996). For the current purposes, this distinction is unimportant (but see Colby and Goldberg 1999; Snyder et al. 1997). This report

demonstrates that a simple 3-layer network is capable of encoding remembered target locations in one of 2 frames of reference. The accuracy and precision of the network's output matches the pattern of behavior seen in animals trained to perform a similar task. The network can track targets when gaze shifts are presented as position, displacement, or velocity signals. When trained with both position and velocity inputs, the model relied more heavily on velocity than position to perform the task. Only networks that used position displayed responses that were modulated by gaze position (gaze position gain fields) in the hidden layer units.

Performance

Posterior parietal neurons, specifically those in LIP, have been shown to encode remembered spatial locations that are likely targets of an impending eye movement (Gnadt and Andersen 1995; Platt and Glimcher 1997; Snyder et al. 1997). These neurons use extraretinal signals to update spatial information to compensate for self-motion (Duhamel et al. 1992; Mazzoni et al. 1996; Powell and Goldberg 1997). However, the appropriate combination of retinal and extraretinal signals depends on the reference frame of the remembered target. Most studies have used targets that remained fixed in the world. Features of stationary (i.e., world-fixed) objects constitute the vast majority of saccade targets; thus one could imagine that the responses observed in LIP neurons are specialized for encoding saccadic targets in a world-fixed frame of reference. However, humans and animals are also capable of directing saccades to gaze-fixed targets (Baker et al. 2003; Israel et al. 1999). Are the same neurons that perform spatial updating capable of suppressing their responses to self-motion signals in the gaze-fixed context? Or does the capacity to encode targets in a reference frame require a network of neurons specialized for encoding targets in that frame?

The model we describe reveals that a simple distributed network of neuron-like units in the hidden layer are capable of flexibly representing targets of saccades in either a world-fixed or gaze-fixed reference frame. Based on previous successful comparisons between hidden layer units and posterior parietal neurons (Xing and Andersen 2000a,b; Zipser and Andersen 1988), we predict that neurons in LIP should also be capable of such flexible responses.

Droulez and Berthoz (1991) demonstrated that world-fixed targets could be tracked entirely within a retinotopic coordinate system, without an explicit world-centered representation. We have further demonstrated that a simple neural network can track targets in both world-fixed and gaze-fixed frames of reference in an oculocentric coordinate system. The network effectively ignores gaze shift information when presented with a gaze-fixed target (Fig. 8), extending the finding of Droulez and Berthoz to show that not only is an explicit world-centered representation not required, but also that updating within the retinotopic coordinate system need not be obligatory.

Model versus behavior

In comparing our artificial neural network to real networks in the brain, it is important for the model to be validated by comparisons of network output to either neuronal or behavioral output. Output is less precise in the world-fixed task than in the gaze-fixed task for both animal (behavior: Baker et al. 2003; neurons: Baker et al. 2002) and model when updating occurs in response to a slow gaze perturbation. In addition, the neural network and the monkeys showed similar edge effects (Fig. 5). Both were hypometric in directing saccades to the most peripheral targets. At the same time, world-fixed saccades to the periphery were less variable than world-fixed saccades to central targets. The fact that the model reproduced aspects of the animals' behavior that were not explicitly reinforced by our training is evidence that the choice of model architecture and parameters was appropriate.

The results from these simulations help to illuminate behavioral data from the monkey. In vivo saccade variability is greater in world-fixed than in gaze-fixed trials. This same pattern is observed from the network when the only source of noise is the gaze signal. These data support the hypothesis that the pattern of behavioral variability reflects, in part, whether a noisy input signal is incorporated into or excluded from the output (Baker et al. 2003). It is also possible that noise is exacerbated on world-fixed but not gaze-fixed trials as a result of imprecise computation by neurons (Shadlen and Newsome 1998).

Although the network serves as a good model of animal behavior in the rotation and pursuit trials, the model did not mimic the behavior observed in saccade trials. On saccade trials, both gaze-fixed and world-fixed memory saccades were equally precise and more accurate than in rotation and pursuit (Baker et al. 2003). Simply shortening the duration of the gaze shift or using a displacement signal as an input had no effect on the pattern of model output. The differences between saccadic and slow gaze shifts may be attributable to the different nature of the movements. At the initiation of a movement, the endpoint is known for a saccade, but not for a pursuit movement or whole-body rotation. In addition, the brain may be optimized for scanning the visual environment using saccades (Niemeier et al. 2003). These differences may result in different mechanisms for integrating self-motion information into the stored representation of salient targets for slow versus saccadic gaze perturbations.

Hidden units share both transformations

When the network is trained, the output layer correctly codes the location of a remembered target in one of 2 reference frames. The activity of the output layer is created by projections from a single population of hidden units that contributes to both the gaze-fixed and world-fixed transformations (Fig. 7). The alternative, that 2 subpopulations of hidden units would emerge (each responsible for one of the transformations), would imply that flexible output requires gating one of 2 dedicated subpopulations of neurons. Instead, we found that a single population of units can encode the locations of remembered targets on both gaze-fixed and world-fixed trials (Fig. 7). We hypothesize that the neurons involved in updating can also flexibly track targets in both tasks.

The model prefers velocity

The neural network model prefers velocity information to perform the updating task. This is supported by improved network performance with velocity compared with position

inputs (Fig. 8*A*) and by the preference for velocity exhibited by a network trained with both position and velocity inputs (Fig. 8*B*). The noise analysis shows that the choice of gaze velocity over gaze position is rather insensitive to the relative signal-to-noise ratios of the 2 inputs (Fig. 8*E*). These results indicate that velocity information may be simpler to incorporate than position information into a dynamically updating network. This may occur because the recurrent layer already functions as an integrator to maintain the memory of the transiently presented target. Transient velocity signals express the shift in gaze, which can be integrated by the network to compensate for gaze shifts. The position input, in contrast, does not directly encode shift in gaze. Instead, the gaze signal must be first differentiated to obtain the change in gaze required for updating. Alternatively, a representation of the original eye position could be stored in memory buffer, and this stored signal subtracted from the final gaze position at the end of the perturbation (Xing and Andersen 2000a).

Gaze displacement signals are intermediate: although they directly encode the magnitude of the gaze shift, they must also be integrated, given that the network must store the gaze shift information after the displacement signal is reset. Networks did not show a preference between velocity and displacement or between displacement and position (Fig. 8, *C* and *D*).

Neural circuits involved in controlling saccades at the level of the superior colliculus and brain stem, like the cortical circuits involved in updating, also use information about on-going gaze shifts. Initially, these low-level control circuits were modeled using only gaze position signals (Jürgens et al. 1981; Robinson 1975) or only gaze velocity (Scudder 1998). However, more recent work suggests that both position and velocity signals are required (Arai et al. 1999) to explain the behavior observed when omnipause neurons are stimulated during a saccade (Keller et al. 1996). By analogy, it would seem likely that the cortex would also use both position and velocity signals.

Explicit gaze position or velocity signals have not been observed in LIP (Bremmer et al. 1997; Fukushima 1997). However, there are many potential sources of these signals in the brain. Horizontal and vertical eye position information can be found in prepositus hypoglossi, the interstitial nucleus of Cajal, and the vestibular nuclei, for example (reviewed in Sparks 2002). Gaze velocity signals resulting from smooth pursuit and from head rotation occur in gaze velocity Purkinje cells in the cerebellar flocculus, the dorsolateral pons, and vestibular nuclei. Smooth pursuit eye movement signals can also be found in a region next to the frontal eye fields (Tian and Lynch 1996) and in area MST (Churchland and Lisberger 2002; Newsome et al. 1988). Gain fields in LIP encode both eye position in the head and head direction on the body (Snyder et al. 1998). Gain fields in parietal area 7a encode gaze position in the world (Snyder et al. 1998). Head direction in the world is commonly found in parietal and perihippocampal head-direction neurons in rats, but in primates the direct encoding of head direction in the world is less clear (Froehler and Duffy 2002; Ono and Nishijo 1999; Rolls 1999). To summarize, neurons can be found in the primate that encode gaze position with respect to the head and body and that encode gaze velocity with respect to the head, body, and world. Whether position in the world is directly encoded in the brain remains an open question. All of the signals described can provide the necessary

gaze perturbation information (gaze position or velocity) for updating.

Gaze position gain fields

A previous model has demonstrated that a distributed neural network can effectively update a memorized spatial location based on a changes in a position signal (Xing and Andersen 2000a). In this network, gain fields for position were observed in the hidden layer. It has been proposed that gain fields may serve as a mechanism for updating, given that gain fields have been observed in many of the areas where spatial signals are updated.

Spatial memories in LIP can also be updated when the whole body is passively rotated and gaze rotates with the body (Baker et al. 2002; Powell and Goldberg 1997). However, a gain modulation by position in the world has not been observed (Snyder et al. 1998). Neurons in LIP display gain fields for head orientation (angular position) on the body, but not for absolute head orientation in the world (Fig. 9*B*). The absence of LIP gain fields for position in the world suggests that an allocentric (world-referenced) position signal is not available to neurons in this area.

How can neurons in LIP update target locations in response to passive whole body rotation without an appropriate gaze position signal? The model we describe indicates that a distributed neural network can use a gaze velocity or displacement signal for this purpose. Such networks do not display gain fields for gaze position, indicating that units in the hidden layer are capable of integrating the velocity signal or storing the displacement signal to generate the appropriately updated target position.

Zipser and Andersen (1988) demonstrated gaze position gain fields in hidden layer units that combine retinal and extraretinal information. Here, we show that gain fields are present only when the extraretinal information is in the form of a gaze position signal. When gaze velocity or gaze displacement is provided instead of gaze position, the network is still able to update the retinal input, but the hidden layer units do not show gaze position gain fields. Thus gain fields are not an inevitable consequence of combining retinal and extraretinal information, but instead may be a signature of the use of position information.

REFERENCES

Andersen RA, Brotchie PR, and Mazzoni P. Evidence for the lateral intraparietal area as the parietal eye field. *Curr Opin Neurobiol* 2: 840–846, 1992.

Andersen RA and Mountcastle VB. The influence of the angle of gaze upon the excitability of the light-sensitive neurons of the posterior parietal cortex. *J Neurosci* 3: 532–548, 1983.

Arai K, Das S, Keller EL, and Aiyoshi E. A distributed model of the saccadic system: simulations of temporally perturbed saccades using position and velocity feedback. *Neural Networks* 12: 1359–1375, 1999.

Baker JT, Harper T, and Snyder LH. Spatial memory following shifts of gaze. I. Saccades to memorized world-fixed and gaze-fixed targets. *J Neurophysiol* 89: 2564–2576, 2003.

Baker JT, White RL, and Snyder LH. Reference frames and spatial memory operations: area LIP and saccade behavior. *Soc Neurosci Abstr* 57: 16, 2002.

Batista AP and Andersen RA. The parietal reach region codes the next planned movement in a sequential reach task. *J Neurophysiol* 85: 539–544, 2001.

Bremmer F, Distler C, and Hoffmann KP. Eye position effects in monkey cortex. II. Pursuit- and fixation-related activity in posterior parietal areas LIP and 7A. *J Neurophysiol* 77: 962–977, 1997.

Britten KH, Shadlen MN, Newsome WT, and Movshon JA. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci* 12: 4745–4765, 1992.

Brotchie PR, Andersen RA, Snyder LH, and Goodman SJ. Head position signals used by parietal neurons to encode locations of visual stimuli. *Nature* 375: 232–235, 1995.

Churchland AK and Lisberger SG. Eye velocity tuning of extraretinal responses in MST. *Soc Neurosci Abstr* 56: 2, 2002.

Colby CL, Duhamel JR, and Goldberg ME. Oculocentric spatial representation in parietal cortex. *Cereb Cortex* 5: 470–481, 1995.

Colby CL, Duhamel JR, and Goldberg ME. Visual, presaccadic, and cognitive activation of single neurons in monkey lateral intraparietal area. *J Neurophysiol* 76: 2841–2852, 1996.

Colby CL and Goldberg ME. Space and attention in parietal cortex. *Annu Rev Neurosci* 22: 319–349, 1999.

Droulez J and Berthoz A. A neural network model of sensoritopic maps with predictive short-term memory properties. *Proc Natl Acad Sci USA* 88: 9653–9657, 1991.

Duhamel JR, Bremmer F, BenHamed S, and Graf W. Spatial invariance of visual receptive fields in parietal cortex neurons. *Nature* 389: 845–848, 1997.

Duhamel JR, Colby CL, and Goldberg ME. The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255: 90–92, 1992.

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

Fukushima K. Corticovestibular interactions: anatomy, electrophysiology, and functional considerations. *Exp Brain Res* 117: 1–16, 1997.

Funahashi S, Bruce CJ, and Goldman-Rakic PS. Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61: 331–349, 1989.

Funahashi S, Bruce CJ, and Goldman-Rakic PS. Visuospatial coding in primate prefrontal neurons revealed by oculomotor paradigms. *J Neurophysiol* 63: 814–831, 1990.

Galletti C, Battaglini PP, and Fattori P. Parietal neurons encoding spatial locations in craniotopic coordinates. *Exp Brain Res* 96: 221–229, 1993.

Gnadt JW and Andersen RA. Memory related motor planning activity in posterior parietal cortex of macaque. *Exp Brain Res* 70: 216–220, 1988.

Goldberg ME and Bruce CJ. Primate frontal eye fields. III. Maintenance of a spatially accurate saccade signal. *J Neurophysiol* 64: 489–508, 1990.

Israel I, Ventre-Dominey J, and Denise P. Vestibular information contributes to update retinotopic maps. *Neuroreport* 10: 3479–3483, 1999.

Jurgens R, Becker W, and Kornhuber HH. Natural and drug-induced variations of velocity and duration of human saccadic eye movements: evidence for a control of the neural pulse generator by local feedback. *Biol Cybern* 39: 87–96, 1981.

Keller EL, Gandhi NJ, and Shieh JM. Endpoint accuracy in saccades interrupted by stimulation in the omnipause region in monkey. *Vis Neurosci* 13: 1059–1067, 1996.

Mainen ZF and Sejnowski TJ. Reliability of spike timing in neocortical neurons. *Science* 268: 1503–1506, 1995.

Mazzoni P, Bracewell RM, Barash S, and Andersen RA. Motor intention activity in the macaque's lateral intraparietal area. I. Dissociation of motor plan from sensory memory. *J Neurophysiol* 76: 1439–1456, 1996.

Mitchell J and Zipser D. A model of visual-spatial memory across saccades. *Vision Res* 41: 1575–1592, 2001.

Mitchell M. *An Introduction to Genetic Algorithms* (*Complex Adaptive Systems*). Cambridge, MA: MIT Press, 1996, p. 166–170.

Nakamura K and Colby CL. Visual, saccade-related, and cognitive activation of single neurons in monkey extrastriate area V3A. *J Neurophysiol* 84: 677–692, 2000.

Nakamura K and Colby CL. Updating of the visual representation in monkey striate and extrastriate cortex during saccades. *Proc Natl Acad Sci USA* 99: 4026–4031, 2002.

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.

Niemeier M, Crawford JD, and Tweed DB. Optimal transsaccadic integration explains distorted spatial perception. *Nature* 422: 76–80, 2003.

Olson CR. Brain representation of object-centered space in monkeys and humans. *Annu Rev Neurosci* 26: 331–354, 2003.

Ono T and Nishijo H. Active spatial information processing in the septo-hippocampal system. *Hippocampus* 9: 458–466, 1999.

Platt ML and Glimcher PW. Responses of intraparietal neurons to saccadic targets and visual distractors. *J Neurophysiol* 78: 1574–1589, 1997.

Powell KD and Goldberg ME. Remapping of visual responses in primate parietal cortex during smooth changes in gaze. *Soc Neurosci Abstr* 23: 17, 1997.

Robinson DA. Oculomotor control signals. In: *Basic Mechanisms of Ocular Motility and Their Clinical Implication*, edited by Lennerstrand G and Bach-y-Rita P. Oxford, UK: Pergamon Press, 1975, p. 337–374.

Rolls ET. Spatial view cells and the representation of place in the primate hippocampus. *Hippocampus* 9: 467–480, 1999.

Rumelhart DE, Hinton GE, and Williams R. Learning internal representations by error propagation. In: *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, edited by Rumelhart DE and McClelland JL. Cambridge, MA: MIT Press, 1986, vol. 1, p. 316–362.

Scudder CA. A new local feedback model of the saccadic burst generator. *J Neurophysiol* 9: 1455–1475, 1988.

Shadlen MN and Newsome WT. The variable discharge of cortical neurons: implications for connectivity, computation, and information coding. *J Neurosci* 18: 3870–3896, 1998.

Snyder LH, Batista AP, and Andersen RA. Coding of intention in the posterior parietal cortex. *Nature* 386: 167–170, 1997.

Snyder LH, Batista AP, and Andersen RA. Intention-related activity in the posterior parietal cortex: a review. *Vision Res* 40: 1433–1441, 2000.

Snyder LH, Grieve KL, Brotchie P, and Andersen RA. Separate body- and world-referenced representations of space in parietal cortex. *Nature* 394: 887–891, 1998.

Sparks DL. The brainstem control of saccadic eye movements. *Nat Rev Neurosci* 3: 952–964, 2002.

Tian JR and Lynch JC. Functionally defined smooth and saccadic eye movement subregions in the frontal eye field of Cebus monkeys. *J Neurophysiol* 76: 2740–2753, 1996.

Walker MF, Fitzgibbon EJ, and Goldberg ME. Neurons in the monkey superior colliculus predict the visual result of impending saccadic eye movements. *J Neurophysiol* 73: 1988–2003, 1995.

Williams RJ and Zipser D. Gradient-based learning algorithms for recurrent neural networks. In: *Back-Propagation: Theory, Architecture and Applications*, edited by Chauvin Y and Rumelhart DE. Hillsdale, NJ: Erlbaum, 1995, p. 433–486.

Xing J and Andersen RA. Memory activity of LIP neurons for sequential eye movements simulated with neural networks. *J Neurophysiol* 84: 651–665, 2000a.

Xing J and Andersen RA. Models of the posterior parietal cortex which perform multimodal integration and represent space in several coordinate frames. *J Cogn Neurosci* 12: 601–614, 2000b.

Zipser D. Identification models of the nervous system. *Neuroscience* 47: 853–862, 1992.

Zipser D and Andersen RA. A back-propagation programmed network that simulates response properties of a subset of posterior parietal neurons. *Nature* 331: 679–684, 1988.