

## A neural network model of sensoritopic maps with predictive short-term memory properties

J. DROULEZ AND A. BERTHOZ

Laboratoire de Physiologie Neurosensorielle du Centre National de la Recherche Scientifique, 15, Rue de l'École de Médecine, 75006 Paris, France

Communicated by Jean-Pierre Changeux, February 21, 1991

**ABSTRACT** Coordinated orienting movements can be accurately performed without direct sensory control. Ocular saccades, for instance, have been shown to be reprogrammed after target disappearance when an intervening eye movement is electrically triggered before the saccade onset. Saccadic eye movements can also be executed toward memorized targets, even when the subject has been passively moved in darkness. Two hypotheses have been proposed to account for this goal-invariance property: either (i) the goal is reconstructed and memorized in the stable frame of reference linked to the environment ("allocentric coordinates") or (ii) the goal is selected and memorized in the sensors-related maps ("egocentric coordinates") and is continuously updated by efferent copies of the motor commands. In this paper, we shall describe a formal neural network based on this second hypothesis. The results of the simulation show that target position can be memorized and accurately updated in a topologically ordered map, using a velocity-signal feedback. Moreover, this network has been submitted to a simple learning procedure by using the intermittent visual recurring afferent signal as the teaching signal. A similar mechanism could be involved in control of limb movement.

Among the various types of orienting movements, visually triggered oculomotor saccades have been extensively studied in the past 2 decades. The stereotyped dynamic characteristics of saccades initially led to the consideration of saccades as an example of ballistic movements elicited from retinal stimulation by a "built-in temporal-pattern generator" (1). In 1975, Robinson (2) argued that visual-target position first has to be recomputed in a craniotopic frame of reference by adding an eye-position efferent copy. His model predicted that when the eyes were displaced by an internal command after the target had disappeared and before the saccade started, the saccade would still end where the target had been. This prediction was confirmed later by Mays and Sparks (3, 4) in monkey and by Viviani and Velay (5) in humans. Tweed and Vilis (6) extended the Robinson model to the three-dimension eye-rotation space by using the formalism of quaternionic algebra to code the desired and instantaneous eye position. Moreover, Guitton *et al.* (7) proposed that target position and gaze position are reconstructed in space coordinates for the programming of eye/head coordinated movements.

However, despite intensive investigations in the various neural structures involved in saccade generation, the neural coding of desired eye position in the orbit (and of target in craniotopic coordinates) has never been found. Becker and Jürgens (8) proposed a modification of the Robinson model, in which a dynamic motor-error signal is computed as the difference between the retinotopic target position and the instantaneous eye displacement, provided by an integrator that can be reset. Waitzman *et al.* (9) have recorded single-

cell activities in the intermediate layers of monkey superior colliculus that are correlated with the dynamic motor error. A correlation has been reported between tecto-reticulospinal neurons discharge and tangential eye velocity in the cat (10) and in the monkey (11).

The use of an integrator that can be reset allows the control of saccade execution without computing the target position in craniotopic coordinates. However, this modification left unsolved the problem of how spatially invariant orienting movements could be specified and memorized: in the double-saccade paradigm, for instance, the dynamic motor error is zeroed after the first saccade (because of the integrator reset), and the second saccade is determined by the stored retinotopic position of the second target. To solve this problem, we assumed that a retinotopic memory map integrates the target displacement that can be predicted from the eye movement in the absence of any visual afferent signal. Such a model also provides an estimate of the dynamic motor error, but it does not require the existence of an integrator that can be reset: the prediction is a continuous mechanism that can be used as a complement or a substitute for the sensory signal. This concept has been proposed in its general context recently (12).

### BASIC CONCEPTS

Only indirect experimental evidence supports the hypothesis of a spatiotopic representation of targets: in the monkey parietal cortex, Andersen *et al.* (13) found neurons that discharge as a nonlinear (multiplicative) function of visual stimulation and eye position in the orbit; although their receptive fields were retinotopically organized, these neurons were interpreted as an intermediate computational step for a craniotopic coding of target position (14). On the contrary, most neuronal structures, known to be directly involved in the programming of visually triggered saccades, exhibit a retinotopic organization. A second hypothesis is that goal invariance is enabled by the ability of the brain to predict any change in goal representation, based on a memory of targets in sensory-based frames of reference, providing that such change could be inferred from other information.

The idea that the brain works as a predictor is a very general and widely accepted notion. Pellionisz and Llinas (15) attributed a predictive function to the cerebellum for sensorimotor coordination: the extrapolation of future target position as a linear combination of estimated target position and velocity constitutes a simple example of such predictive function. We have developed a model that postulates that sensory and motor information are combined according to the constraint of maximal internal coherence between central representations (16). The same principle could be applied to the prediction, based on central estimates of eye and target position and velocity, of where a visual target should be located in a retinotopic map.

Let  $f(x, y, t)$  equal the neuronal activity of a set of neurons within a retinotopic map at time  $t$ . The choice of the  $x, y$  retinal coordinates is arbitrary. Each neuron is identified by

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

the retinotopic position of its receptive field. Let  $(D_x, D_y)$  equal the target-displacement vector with respect to the retina from time  $t$  to time  $t + D_t$ . If the retinotopic map correctly predicts the new target position, activity in the map should be shifted accordingly:

$$f(x + D_x, y + D_y, t + D_t) = f(x, y, t). \quad [1]$$

The first-order Taylor development of Eq. 1 yields the following equation:

$$f(x, y, t + D_t) = f(x, y, t) - D_x \cdot df/dx(x, y, t) - D_y \cdot df/dy(x, y, t). \quad [2]$$

If the target were fixed in space, its displacement relative to the retina would be only due to eye velocity. Let us note  $V_x$  and  $V_y$ , the two components of eye velocity, expressed in retinal coordinates. The target-displacement vector relative to the retina is directly obtained from the eye-velocity components:

$$D_x = -D_t \cdot V_x \text{ and } D_y = -D_t \cdot V_y. \quad [3]$$

Replacing these terms in Eq. 2 yields the following equation:

$$f(x, y, t + D_t) = f(x, y, t) + D_t(V_x \cdot df/dx + V_y \cdot df/dy). \quad [4]$$

The temporal variation of the activity of a given neuron in the map is proportional to the dot product of the eye velocity and the gradient vector of activity. The temporal step value  $D_t$  is related to the transmission and synaptic delays in the neuronal loops. The spatial step value implies that the activity-gradient vector  $(df/dx, df/dy)$  is approximated by the finite difference between activities of neighboring neurons. The overlapping of adjacent receptive fields allows generalization of the discrete version of Eq. 4 as follows:

$$f(x_i, y_i, t + D_t) = \sum_j (a_{ij} + D_t \cdot V_x \cdot b_{ij} + D_t \cdot V_y \cdot c_{ij}) \cdot f(x_i + dx_j, y_i + dy_j, t). \quad [5]$$

The coefficients  $a_{ij}$  expressed that the activity of a given neuron could be approximated by a linear interpolation of neighboring activities, whereas the coefficients  $b_{ij}$  and  $c_{ij}$  expressed that the activity gradient is estimated by an anti-symmetric linear combination of neighboring activities.

Quantitative simulations have been done on regular maps (one or two dimensions) organized in modules or pixels, equally distant from each other. The two-dimensional map was formed of  $31 \times 31$  modules of four neurons, whereas the one-dimensional map was formed of 50 modules of three neurons. Three types of neurons were specified (Fig. 1): The input neuron (type I) receives weighted inputs from the main neurons (type M) and interneurons (type P) of the neighboring modules. Each input neuron also receives an external excitatory input  $S$  (visual stimulation). When visual information is available, the input neuron activity reflects the error between the stored information and the visual signal; the synaptic weights of the M and P projections are adjusted to continuously minimize this error. In the absence of visual stimulation, activity of the type I neuron is then a mirror image of the predicted activity in the visual pathway. The main output neuron M compares the external input  $S$  with the activity of the corresponding input neuron; the resulting activity  $f(x, y, t)$  is a nonlinear sigmoid function of the difference between the two inputs. When the visual information is available, activity in the type I neuron is minimized, and the main neuron just follows visual input  $S$ . In the absence of visual information, the main neurons are inhibited by activity of the corresponding type I neurons, except when this activity is low—that is, near the predicted target position. The interneurons P compute the product of M activity and one component of the eye-velocity signal. There are, at least, two type P interneurons in a two-dimensional map, one for each eye-velocity component.

The network connectivity is local in the sense that each module receives inputs only from neighboring modules; the spatial extent of this local connectivity is not a critical factor, but it must be greater than the size of the receptive field. The local connectivity implements Eq. 5 as follows:  $a_{ij}$  represents

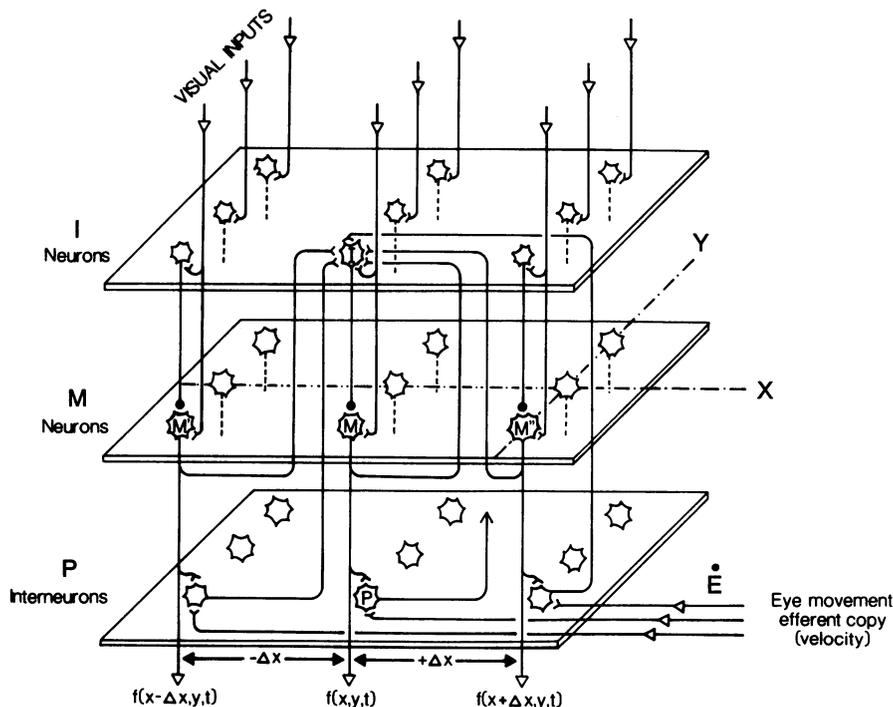


FIG. 1. Neuronal architecture of the dynamic memory map (see text). The three layers correspond to the three types of neurons: input neurons (I), main neurons (M), and multiplier interneurons (P).

the weight of the projection to type I neuron located at coordinates  $(x_i, y_i)$  in the map, originated from a type M neuron located at  $(x_j, y_j)$ ; in regular maps, the differences  $x_j - x_i$  and  $y_j - y_i$  are multiples of the size of 1 pixel;  $b_{ij}$  (respectively,  $c_{ij}$ ) represents the weight of the input to the type I neuron located at coordinates  $(x_i, y_i)$  in the map, originating from the type P neuron located at  $(x_j, y_j)$ , which receives the eye-velocity component  $V_x$  (respectively,  $V_y$ ).

The temporal step was set to 5 msec; coordinates and distances were evaluated in pixels. Fig. 2 shows the variation of synaptic weights  $a_{ij}$  and  $b_{ij}$  as a function of the distance between source and target neurons.

A learning procedure was applied only on the one-dimensional map. The synaptic weights were initially set to zero. Visual stimulations (Gaussian-like function) were presented to the type I and M neurons. A predefined sequence of fixations and saccades was then applied. During fixation (duration, 40 steps—i.e., 200 msec), the visual input was kept constant over the time, and the eye-velocity signal  $V_x$  was set to zero. During saccade (duration, 10 steps—i.e., 50 msec), the eye-velocity signal, randomly chosen in the range  $-400, +400$  pixels/sec, was input to the type P neurons, and the mean position of the visual input was displaced accordingly—that is, in the opposite direction and with the same absolute velocity. A modified version of the Hebbian learning rule (17) was applied to the synaptic weights  $a_{ij}$  and  $b_{ij}$ , so that the weighted sum of reverberating activities predicts, as accurately as possible, the visual input at the next step. The synaptic weight of a given input is changed by an amount

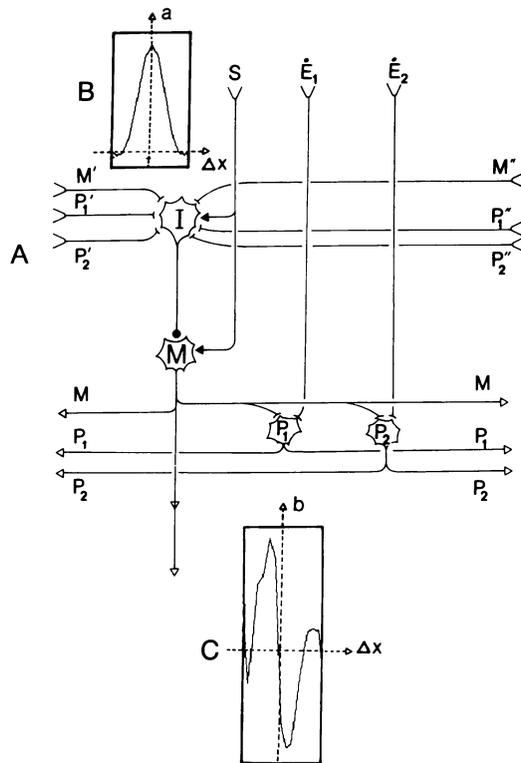


FIG. 2. Detailed description of the modular architecture of the dynamic memory map. (A) Only one module containing three types of neurons is shown. E, eye movement efferent copy (velocity); S, external excitatory visual input (see text). (B and C) Synaptic weights, obtained after the learning phase, as a function of the coordinate difference between the target and the source neuron. (B) Synaptic weights of projections from neighboring M neurons. Note the symmetric shape and existence of a small lateral inhibition. (C) Synaptic weights of projections from neighboring P neurons. Note the expected antisymmetric shape (computation of the activity gradient).

proportional to the product of that input activity and an error signal to be minimized. In our model, this error signal is directly related to the type I neuron activity (that is, the difference between the visual input and the weighted sum of reverberating activities originating from type M and type P neuronal activities). Therefore, a given synaptic weight increases (respectively, decreases) when the presynaptic activity and the postsynaptic potential are negatively (respectively, positively) correlated. After the learning phase was completed, several tests were done in the absence of visual afferent signal, and the activity of type M neurons  $f(x, y, t)$  was followed during fixations and saccades. No learning procedure was applied to the two-dimensional map. The synaptic weights  $a_{ij}$  were defined as a Gaussian function of the distance between interconnected modules (radius, 2 pixels), and the synaptic weights  $b_{ij}$  (respectively,  $c_{ij}$ ) were proportional to the first spatial derivative of this function along the  $x$  axis (respectively, the  $y$  axis). The input-output function of each neuron was a sigmoid function, to which a low-pass filter was applied (time constant, 5 msec). The activity in the map was initialized or reset by superimposing on the stored activity a Gaussian-like stimulation  $S$  (radius, 5 pixels).

RESULTS

**Storage of Information.** The stability of the stored information is due mainly to the nonlinear, sigmoid-shaped input-output characteristic of the type M neuron. Fig. 3 shows the distributed activity in the map, just after the visual input (Fig. 3B) and after 10,000 steps—that is, 50 sec (Fig. 3C) in the one-dimensional map. The shape of the distributed activity slightly changes at the beginning of the test in the two-dimensional map: this is due to the fact that synaptic weights in the two-dimensional map are not precisely adjusted to the size of the receptive fields, as they are after the learning phase in the one-dimensional map. Nevertheless, the activity tends toward a stable configuration. Note also that the centroid of

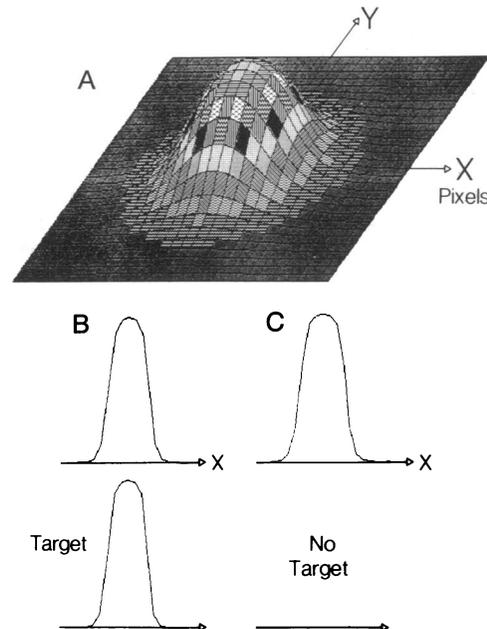


FIG. 3. Stability of stored activity in the map. (A) Stored activity in the bidimensional map. (B and C) Stored activity in the one-dimensional map (upper curve), as compared with the activity in the visual efferent signal (lower curve) during target presentation (B) and 50 sec after target disappearance (C). There is a very small modification in the shape of stored activity between 50 and 500 msec.

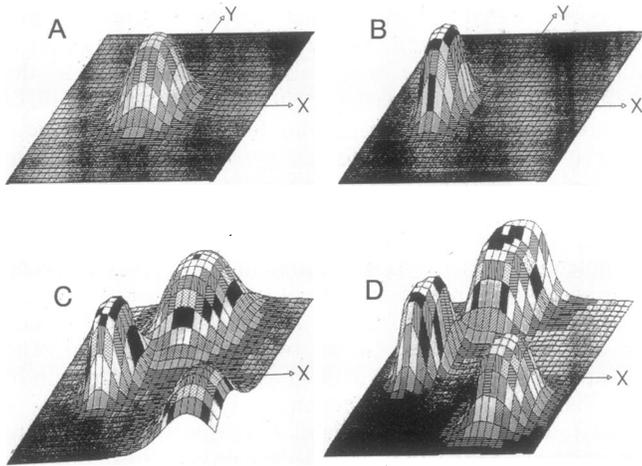


FIG. 4. Example of the shifting of activity in the bidimensional map during saccades. (A) Initial activity (single target). (B) Final activity after 50 msec (horizontal saccade). (C) Initial activity (three targets). (D) Final activity after an oblique saccade. Distances between peaks of activity are preserved in spite of the small deformation of the overall shape.

the distributed activity, which is supposed to code the target position, is maintained nearly unchanged.

The use of a spatial code allows the simultaneous tracking of a great amount of information (several targets can be simultaneously stored in the same map). This advantage is paid for by a loss of precision as compared with a frequency code. The crucial parameter here is the overlapping factor—that is, the ratio of the receptive field size versus the distance between adjacent modules. The number of targets that can be stored decreases, and the precision roughly increases in proportion to the overlapping factor. We have tested this effect on the one-dimensional map. When the two targets are sufficiently far from each other, they elicit two well-separated mountains of activity. When the two targets are close enough, the two mountains slightly overlap. Such configuration is not stable, and after several steps, it evolves into a unique mountain that encompasses the two initial peaks. The model predicts that when several neighboring targets are presented simultaneously, the elicited saccade will end on their centroid.

**Displacement of Activity.** Fig. 4 shows that activity in the map is shifted without noticeable change in amplitude and overall distribution when an eye-velocity command is delivered to the type P neurons. At the end of the eye movement, the new distribution of activity is kept stationary, because of the short-term memory property of the map described above. The displacement between two successive steps is nearly equal and opposite to the eye displacement, so that the

distributed activity follows the kinematic characteristic of the eye movement. We have performed several simulations on the one-dimensional map to test whether there is an optimal eye velocity beyond which the map could not work correctly. Fig. 5 displays the error on the estimated target displacement as a function of saccade amplitude for various saccade durations. The results show that this error is minimized for a given saccadic velocity ( $\approx 230$  pixels per sec) but remains small for lower saccadic velocity.

**Simulation of the Mays and Sparks Experiment.** At the beginning of the simulation, the target was in position T and has induced a mountain of activity in the upper region of the map. The electrical stimulation has elicited another peak of activity (S). During the first saccade, the eyes moved towards S—that is, horizontally and to the right. The corollary discharge of eye velocity was input to the dynamic memory (on layer P) and induced a left horizontal shift of both mountains of activity, of an amount equal to the amplitude of the saccade. At the end of the first saccade, the “stimulation mountain” (S’) is located on the fovea and the “target mountain” in the upper left quadrant (T’). Therefore, when the second saccade is triggered, it is directed towards T’—that is, obliquely to the up and left quadrant. From the information stored in the predictive dynamic memory map, the amplitude and direction of the second saccade are equal to those obtained were the target still visible. Simulations of orienting movements to memorized locations in space can also be done on the same neuronal network. In head-free condition, the signal input to the P neurons must be the gaze velocity (sum of the eye in reference to head and head in reference to space velocities).

Contrary to models assuming a craniotopic representation of target position, the dynamic memory model predicts that the error on the estimated target displacement and, consequently, the dispersion of saccades toward memorized targets should slightly increase with the number and amplitude of intervening movements.

## DISCUSSION

**Sensoritopic Versus Spatiotopic Representation.** We have shown that the spatial invariance of goals can be implemented on predictive sensoritopic maps rather than on spatiotopic maps. The interesting property of this mechanism is that it will give the appearance of a saccade coded in spatiotopic coordinates, although the whole process has been made entirely in retinotopic coordinates. Recordings of neuronal activity in the main structures involved in the generation of orienting movements (superior colliculus, frontal eye field, posterior parietal cortex) in relation to the visual stimulation exhibit a clear retinotopic receptive-field organization, while direct evidence of spatiotopic or even craniotopic represen-

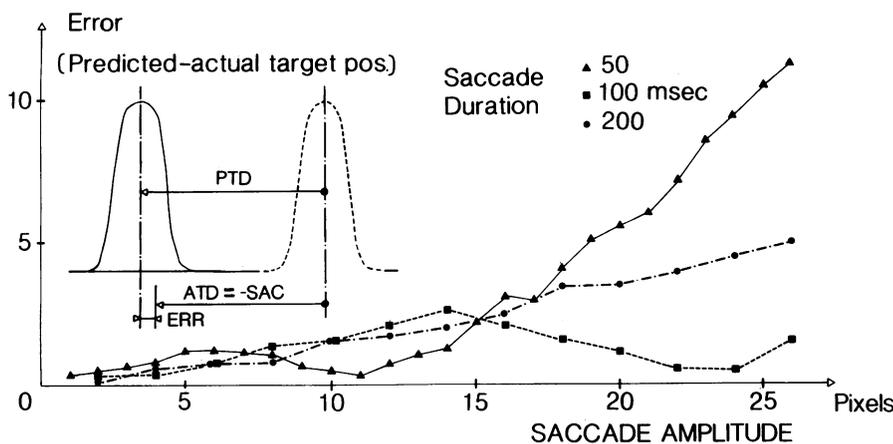


FIG. 5. Prediction error. Difference between predicted target displacement (PTD) and actual target displacement (ATD) in retinal coordinates is plotted as a function of the amplitude of the saccade (in pixels) for various saccadic durations: 200 msec ( $\bullet$ ), 100 msec ( $\blacksquare$ ), and 50 msec ( $\blacktriangle$ ). Minimal errors are obtained with the 50-msec (amplitude, 11 pixels) and 100-msec (amplitude, 23 pixels) saccade, consistent with an optimal displacement at a fixed velocity of  $\approx 230$  pixels per sec.

tation of target position are still lacking. The controversy between possible "mapping" or "updating" mechanisms is, however, still unsolved (18).

Various neuronal behaviors have been described in these structures, with a continuum from purely visual responses to visual-motor and saccade-related responses. Quasi-visual cells and visual-motor cells exhibit a sustained activity when the eye movement is delayed: they are then good candidates to be the substrate of the M units of our model. The receptive fields in the superficial layer and the motor fields in the intermediate and deep layers of collicular neurons are in register and respect the retinal topography. Such anatomical and functional correspondence is less evident in the frontal eye field and in the posterior parietal cortex. However, this correspondence does not restrain the application of our model to the superior colliculus because the connectivity of the network is determined by the functional—and not by the anatomical—organization.

**Velocity or Position Feedback.** In our model, the expected target position relative to the retina is continuously updated by a gaze-velocity signal. It is not clear, however, whether this velocity input is anticipatory (expected gaze velocity) or not (actual gaze velocity), or even whether the eye velocity is input as such or derived from successive eye positions. Neuronal activity in the frontal eye field and in the superior colliculus has been shown to be modulated by saccadic and presaccadic movements. Moreover, Moschovakis *et al.* (19) have shown that the output of primate superior colliculus is fed back to both ipsi- and contralateral superior colliculus, directly to the superficial layer and indirectly to the intermediate and deep layers.

When the eye movement occurs in the light, against a textured background, the gaze-velocity signal could be directly extracted from the retinal slip. During saccades, it is more likely extracted from the oculomotor command itself. Several structures have been shown to have discharge-rate profiles related to eye velocity: excitatory burst neurons in the brain stem, prepositus hypoglossi neurons, etc. However, it is not necessary that the eye velocity signal used to drive the shift of activity is explicitly input to the predictive memory map; it could, instead, be reconstructed within the map, as the difference between two eye position-related signals with different dynamics or as the difference between instantaneous and delayed eye position signals. Firing discharges in the prepositus hypoglossi, for instance, are related, in various proportion, to horizontal eye position and velocity (20). This reconstruction could occur, for instance, at the level of the type P interneurons. In this case, the type P interneuron activity would be exactly the one described by Andersen *et al.* (13) in the posterior parietal cortex of monkeys: they present retinotopic receptive fields to visual stimulation, and their firing discharges are modulated by eye position. This modulation is also required to distribute the desired gaze-velocity command among the various effectors involved in the orienting movements.

**Generalization.** The mechanism described above may be useful to an animal or human to orient with either the eye, the head, or the body, and this mechanism could also be used in reaching. With this mechanism the desired movement is

expressed in a very general manner by using global representations and can be implemented at the periphery by any mechanical effector (eye, head, body, or even arm). Local networks containing neuronal "models" of the mechanical plants could transform this velocity command into a final position. We have further developed the model to include the final oculomotor integrator (21). This integrator operates in local frames of reference. The theory predicts also that there should be local immediate premotor neurons for the independent control of either eye or head movements. When the head is moving, the feedback signals of eye velocity are not sufficient to update the map activity; a head-velocity signal must be available. We know that this signal can be derived from vestibular information or neck proprioception.

This work was supported by a European Economic Community Grant from Esprit II Basic Research Program.

1. Westheimer, G. (1954) *AMA Arch. Ophthalmol.* **52**, 710–724.
2. Robinson, D. A. (1981) in *Models of the Oculomotor Behavior and Control*, ed. Zuber, B. L. (CRC, Boca Raton, FL), pp. 21–41.
3. Mays, L. E. & Sparks, D. L. (1980) *Science* **208**, 1163–1165.
4. Sparks, D. L. & Mays, L. E. (1983) *J. Neurophysiol.* **49**, 45–63.
5. Viviani, P. & Velay, J. L. (1987) in *Eye Movements: From Physiology to Cognition*, eds. O'Regan, K. & Levy Schoen, A. (Elsevier, Amsterdam), pp. 69–78.
6. Tweed, D. & Vilis, T. (1985) *Biol. Cybern.* **52**, 219–227.
7. Guitton, D., Munoz, D. & Galiana, H. (1990) *J. Neurophysiol.* **64** (2), 509–531.
8. Becker, W. & Jurgens, R. (1979) *Vision Res.* **19**, 967–983.
9. Waizman, D., Optican, L. M. & Wurtz, R. E. (1988) *Exp. Brain Res.* **112**, 1–4.
10. Munoz, D. P. & Guitton, D. (1988) *Soc. Neurosci. Abstr.* **13**, 112.9.
11. Rohrer, W. H., White, J. M. & Sparks, D. L. (1987) *Soc. Neurosci. Abstr.* **13**, 1092.
12. Berthoz, A. & Droulez, J. (1991) in *Motor Control: Concepts and Issues*, eds. Humphrey, D. R. & Freund, J. R. (Wiley, Chichester, U.K.).
13. Andersen, R. A., Essick, G. K. & Siegel, R. M. (1985) *Science* **230**, 456–458.
14. Zipser, D. & Andersen, R. A. (1988) *Nature (London)* **331**, 679–684.
15. Pellionisz, A. & Llinas, R. (1979) *Neuroscience* **4**, 323–348.
16. Droulez, J. & Darlot, C. (1989) in *Attention and Performance*, ed. Jeannerod, M. (Lawrence Erlbaum Assoc., Hillsdale, NJ), Vol. 13, pp. 495–526.
17. Rumelhart, D. D., Hinton, G. E. & Williams, R. J. (1989) in *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*, eds. Rumelhart, D. E., McClelland, J. L. & PDP Research Group (Bradford Books/MIT Press, Cambridge, MA), pp. 318–362.
18. Schlag, J. & Schlag-Rey, M. (1990) *Trends Neurosci.* **13** (10), 410–414.
19. Moschovakis, A. B., Karabelas, A. B. & Highstein, S. (1988) *J. Neurophysiol.* **60**, 232–302.
20. Lopez-Barneo, J., Darlot, C., Berthoz, A. & Baker, R. (1982) *J. Neurophysiol.* **47**, 329–352.
21. Droulez, J. & Berthoz, A. (1991) in *The Oculomotor System*, eds. Shimazu, H. & Shinoda, Y. (Elsevier, Amsterdam), in press.