

- Lennard, P. R., and Stein, P. S. G., 1977, Swimming movements elicited by electrical stimulation of turtle spinal cord: I. Low-spinal and intact preparations, *J. Neurophysiol.*, 40:768–778.
- Pearson, K. G., 1993, Common principles of motor control in vertebrates and invertebrates, *Annu. Rev. Neurosci.*, 16:265–297. ♦
- Schöner, G., Jiang, W. Y., and Kelso, J. A. S., 1990, A synergetic theory of quadrupedal gaits and gait transitions, *J. Theoret. Biol.*, 142:359–391.

- Shik, M. L., and Orlovsky, G. N., 1976, Neurophysiology of locomotor automatism, *Phys. Rev.*, 56:465–501.
- Stafford, F. S., and Barnwell, G. M., 1985, Mathematical models of central pattern generators in locomotion: III. Interlimb model for the cat, *J. Motor Behav.*, 17:60–76.
- Taga, G., Yamaguchi, Y., and Shimizu, H., 1991, Self-organized control of bipedal locomotion by neural oscillators in unpredictable environment, *Biol. Cybern.*, 65:147–159.

Gaze Coding in the Posterior Parietal Cortex

Pietro Mazzoni and Richard A. Andersen

Introduction

An issue of active debate in the area of biological motor control is what coordinate frames the nervous system uses to represent sensory stimuli and planned movements. Much evidence suggests that a portion of the posterior parietal cortex (PPC) of the primate brain participates in sensory to motor coordinate transformations. Specific areas within the PPC appear to compute such transformations to program the direction of gaze. We review a set of neural network models developed in our laboratory to study how such coordinate transformations might be achieved by neurons in the PPC.

Neuronal Properties and Presumed Function of the Primate's Posterior Parietal Cortex

Because an animal's sensory and motor organs can move relative to one another, a requirement of sensorimotor integration is the transformation of spatial locations across coordinate frames. Early studies of the monkey's PPC suggested that this area plays a role in the integration of visual perception and motor behavior because neurons were found that responded to visual stimuli and to changes in eye position (reviewed in Andersen, 1987). The portions of the visual field in which luminous stimuli elicited responses—i.e., their receptive fields—corresponded to particular retinal locations. As the monkey looked in different directions, the receptive fields maintained their retinal location but were modulated by eye position. These were called *spatial gain fields* because eye position acted as a gain on the visual response. For most neurons, the modulation had a planar component—i.e., proportional to the horizontal and/or vertical eye position.

The Zipser-Andersen Model

The properties of PPC neurons suggested that individual neurons were unlikely to subserve spatial computations. Being sensitive to both retinal location over a large area and to eye position, a single neuron's activity is an ambiguous signal of stimulus location. This location could in principle be retrieved, however, from the pooled activity of a group of such neurons. Zipser and Andersen developed a neural network to study how an ensemble of neuron-like model units might solve the coordinate transformation problem (Zipser and Andersen, 1988). The aim was to examine the properties of individual units that were trained to solve the problem as a group. If the brain was indeed encoding spatial locations in the distributed pattern of activity of many parietal neurons, then some features of the brain's algorithm might emerge in the model network too.

The Zipser-Andersen network model (Figure 1) is a three-layer feedforward network whose input units carry signals

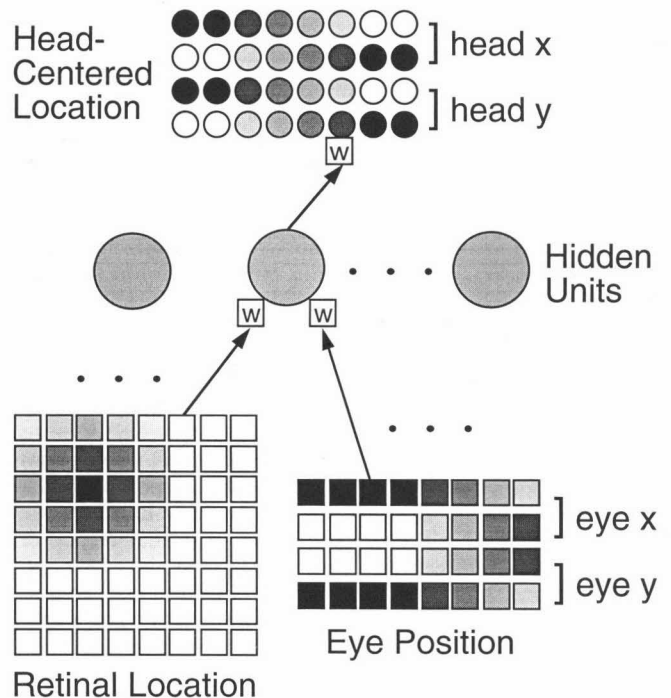


Figure 1. The Zipser-Andersen network model. The input layer consists of 64 units encoding retinal position topographically and 32 units encoding eye position linearly. Each hidden unit has a logistic activation function and projects to all output units. These are 32 logistic units trained to encode the vector sum of the retinal and eye position locations (see Zipser and Andersen, 1988, for details).

known to be available to the PPC. The output layer was trained to encode the head-centered location of the visual stimulus. The task was to perform vector addition of the stimulus's retinal position and the eye position. After the network was trained using backpropagation, its hidden units were found to respond to visual stimuli and to eye position very much like PPC neurons. They had retinotopic visual receptive fields whose activity profiles were modulated by eye position—i.e., spatial gain fields—and the modulation was largely planar. The receptive fields also had shapes similar to those of PPC neurons.

It was thus shown that a layered network can learn to transform retinal coordinates into head-centered ones using the input signals available to the PPC. This result is consistent with the adaptability of spatial behavior. Moreover, the hidden units' representation of spatial information was very similar to

that found in PPC neurons. These neurons can thus play a similar role in the organism—i.e., build up an intermediate representation between input and output stages that is part of the coordinate transformation computation. The network demonstrated explicitly that units with PPC neurons' properties contain a distributed representation of space that is sufficient for accurate localization. Finally, the similarity of response properties suggested that the network and the brain may employ a common strategy in solving the coordinate transformations problem.

Body-Centered Coordinates

Having the problem's solution programmed in a network model made it possible to further investigate what algorithms the PPC may indeed be using through analysis and manipulations of the model. An immediate question was how locations could be encoded in other coordinate frames. The transformation from retinal to head-centered coordinates has a natural application in the programming of eye movements, the eyes having to move to particular positions relative to the head. Large gaze shifts, however, are achieved by coupled movements of the eyes and head; in this case, the target's position must be calculated in body-centered coordinates. Evidence from lesion studies suggests that the PPC is necessary for the proper execution of not only eye movements but other forms of spatial behavior as well. Could the Zipser-Andersen network be modified to compute body-centered coordinates; and if so, what predictions would it make about the PPC?

Goodman and Andersen (1990) added a group of units encoding head position to the input layer of the Zipser-Andersen network and trained this new network to produce body-centered locations at the output layer. The hidden units were found to be sensitive to all three input types. They had retinotopic visual receptive fields modulated by both eye and head position, each in a planar fashion. In other words, they developed planar "gaze fields," that is, linear modulation of visual responses along a particular direction of gaze, which is the sum of eye and head positions. Moreover, the "eye" gain field of a given hidden unit was always aligned (with the same direction and slope) with the same unit's "head" gain field. This was a natural solution for the network given the constraints of its architecture (the eye and head position inputs produced signals in very similar formats) and of the problem (eye and head position are indeed coupled for a given spatial position). The result suggested, however, that if the PPC subserves coordinate transformations beyond the head-centered reference frame and does so with an algorithm analogous to the neural network's strategy, then it should contain units with gaze fields similar to those of the network just described. Such units have recently been identified in the PPC (Brotchie and Andersen, 1991). Brotchie and Andersen trained monkeys to look in various directions by moving their eyes alone or by moving both their eyes and their head. A population of PPC neurons had visual responses modulated equivalently by eye or head position. These gaze fields were largely planar and the direction of eye and head position modulation was the same.

How the Neural Network Transforms Coordinates

Goodman and Andersen (1990) outlined a simple explanation of how the network performs coordinate transformations. Over the course of learning, each hidden unit develops a "preferred direction," that is, a direction in its input space along which to maximally modulate its activity. By maximal modulation, we mean that an input vector parallel to the preferred

direction produces the unit's largest activation, and a vector in the opposite direction produces the smallest activation (or largest inhibition from the resting activity level). The hidden units of the Zipser-Andersen network align their sensitivity in retinal space and in eye position space, and develop an eye position response field that approximates a plane oriented along what becomes the unit's preferred direction (direction \mathbf{a}_i for the i th hidden unit in the network). A hidden unit effectively collapses the multidimensional signal of the retinal and eye position units into two two-dimensional vectors, one for retinal and one for eye position (\mathbf{r} and \mathbf{e} , respectively). The goal is to add these two vectors to obtain the head-centered position vector, \mathbf{h} :

$$\mathbf{r} + \mathbf{e} = \mathbf{h}$$

A hidden unit's activation is proportional to the dot product of its input vectors and its preferred direction (i.e., its input weight vector):

$$\alpha_i \approx \mathbf{r} \cdot \mathbf{a}_i + \mathbf{e} \cdot \mathbf{a}_i$$

Each hidden unit thus extracts the components of the retinal and eye position vectors along its preferred direction and adds them. Because these vectors' components are added at the hidden unit's input, the output of each hidden unit effectively consists of the component of the head-centered vector along the unit's preferred direction. Formally, because

$$\mathbf{r} \cdot \mathbf{a}_i + \mathbf{e} \cdot \mathbf{a}_i = \mathbf{h} \cdot \mathbf{a}_i$$

then

$$\alpha_i \approx \mathbf{h} \cdot \mathbf{a}_i$$

The preferred directions of the hidden units span the two-dimensional input space so that the retinal and eye position vectors are decomposed without losing information. These components are combined again at the output layer to give the vector that is the sum of the retinal and eye position vectors. A single hidden unit's operation can thus be described as a sum of dot products, and is an elegant way of adding two vectors that are encoded in the activity of many input units.

A notable feature of the distributed representation of the Zipser-Andersen network is the absence of topography in the hidden layer. Maintenance of topographic relationships across processing stages can be an effective mechanism for processing spatial information. Several models of saccade generation, for example, use representations with well-defined spatial relationships in order to generate an appropriate saccadic command to look at a sensory stimulus (e.g., Droulez and Berthoz, 1991; Dominey and Arbib, 1992). In these models, the saccade vector is determined by which units are active within a given stage. Units in the hidden layer of the Zipser-Andersen network, on the other hand, are connected to every input and every output unit, and encode the head-centered position vector without regard to any input or output topography. The output vector is determined not by which units are active but by the activity level of every unit in the hidden layer. It is not clear whether PPC areas are topographically organized. The Zipser-Andersen model demonstrates that PPC neurons can transmit to other cortical areas the head-centered position of a stimulus, encoded in their collective firing rate, without any topographic organization.

Perturbing the Model

Stimulation of the Lateral Intra-Parietal (LIP) area in the PPC—a region that directly projects to eye movement centers and that is active during the programming of saccadic eye move-

ments—elicits saccadic eye movements (Shibutani, Sakata, and Hyvärinen, 1984). Goodman and Andersen (1989) simulated the effect of electrically stimulating the PPC by setting the output of a hidden unit to its maximum possible value and interpreting the new position encoded by the output layer as the endpoint of the simulated gaze shift. This process was repeated for many initial eye positions. Because PPC neurons encode spatial locations in neither a retinotopic nor a head-centered reference frame, the effect of varying eye position on the elicited eye movement is not intuitively predictable. Knowing how the network computes coordinate transformations, however, allows us to predict what the effect should be.

Because a hidden unit's activity encodes the component of the head-centered vector along the unit's preferred direction, maximal activation of that unit will shift the network's output along the unit's preferred direction. This direction is encoded in the unit's weights and so should not be affected by the values of the inputs; i.e., the direction of the eye movement will be independent of the starting eye position. The movement's amplitude, on the other hand, depends on how far the unit's initial activation level is from the unit's maximum possible activation. The initial activation level is determined by the unit's input, and thus by the initial eye position. Thus, we expect the amplitude of the simulated saccades to depend on the initial eye position. In particular, as the initial eye position changes along directions orthogonal to the unit's preferred direction, the elicited saccade should be affected very little, because the unit's initial activation will remain the same. As the initial eye position changes along the same direction as the unit's preferred one, on the other hand, the unit's initial activation will vary, and therefore so will the change in the position encoded at the network's output layer when stimulation brings the hidden unit's activation to maximum.

The pattern of eye movements just described was indeed obtained by stimulation of most hidden units in a trained network (Goodman and Andersen, 1989). The elicited saccades had very similar directions from all starting eye positions, but their amplitude decreased as the eye position was shifted along one direction. The direction of this amplitude decrease was very similar to the direction of the elicited eye movement, indicating that the saccades were getting smaller as the eye moved along the unit's preferred direction, as predicted.

Thier and Andersen (1991) found that stimulation of area LIP elicited a pattern of saccades similar to that obtained from the neural network. The saccades evoked from various initial eye positions were all in the same direction. Their amplitude decreased as the starting eye position was moved in one direction, and remained the same as the initial position varied along the orthogonal direction.

Encoding the head-centered location of a stimulus is not the only way in which a saccade to that stimulus can be programmed. Another commonly proposed scheme maps the sensory vector falling on the retina (from the fovea to the stimulus's image) directly into a motor command encoding the required saccade vector, without ever computing the head-centered location of the stimulus. This method still requires some mechanism for keeping track of eye position, so that an appropriate saccade can be made to targets that appeared before one or more intervening eye movements. One such mechanism updates the planned saccade vector based on the last eye movement made. This method has been postulated as a cortical mechanism for saccade planning (Goldberg, Colby, and Duhamel, 1990) and has been used in saccade-generation models. In these models, the future saccade vector is remapped (see DYNAMIC REMAPPING) based either on the integrated eye velocity signal from the intervening saccade (Droulez and

Berthoz, 1991) or on a damped copy of the intervening saccade's eye position signal (Dominey and Arbib, 1992). The Zipser-Andersen model does not address the issue of multiple saccade plans. Extending the model to handle sequences of saccades, however, would not require a remapping scheme that kept track of intervening saccades. All saccade targets would be directly encoded in head-centered coordinates as they appear, and a saccade to each could be planned based only on the current eye position, independently of past eye movements.

Biological Plausibility of the Learning Algorithm

The biological plausibility of the Zipser-Andersen model was an issue of concern because the backpropagation algorithm is an unlikely candidate as a biological learning mechanism. To address this issue, Mazzoni, Andersen, and Jordan (1991) trained a neural network to perform the retinal-to-head-centered coordinate transformation using a reinforcement learning rule developed by Barto and Jordan (1987). This algorithm adjusts the network's connections based on a single error signal computed from the network's overall performance and on the local presynaptic and postsynaptic activation for each connection. Because it combines a reinforcement signal with Hebbian updating of connection strength, it is biologically more plausible than backpropagation (see REINFORCEMENT LEARNING). The hidden units of this network developed gain fields and receptive fields virtually identical to those of the backpropagation-trained networks. The networks' algorithm for computing coordinate transformations, therefore, did not depend on the specific learning mechanism used. The fact that this model learned the computed coordinate transformations based on a simple reinforcement signal also supported the idea that PPC neurons can learn to solve this task from simple feedback signals directly available to the nervous system. Moreover, because the properties of the hidden units are not specific to backpropagation training, the use of backpropagation in the original model does not invalidate its role as a model of PPC function.

Conclusion

The Zipser-Andersen network has been a valuable tool in the study of the PPC. It helped put into an explicit theoretical context many experimental results, and predicted a few additional ones. The original data was not easily summarized by an intuitive coding scheme, partly because the experiments addressed how neurons encode more than one variable. The neural network paradigm provided a framework for developing an intuition about the distributed representation of several variables. As more experiments address the encoding and interactions of several parameters in the nervous system, we expect neural networks to continue to fruitfully assist our investigations of nervous system functions.

Road Map: Primate Motor Control

Background: Perceptrons, Adalines, and Backpropagation

Related Reading: Grasping Movements: Visuomotor Transformations

References

- Andersen, R. A., 1987, The role of the inferior parietal lobule in spatial perception and visual-motor integration, in *The Handbook of Physiology, Section 1: The Nervous System, Vol. IV, Higher Functions of the Brain, Part 2* (F. Plum, V. B. Mountcastle, and S. T. Geiger, Eds.), Bethesda, MD: American Physiological Society, pp. 483–518. ♦
- Barto, A. G., and Jordan, M. I., 1987, Gradient following without backpropagation in layered networks, in *Proceedings of the IEEE*

- International Conference on Neural Networks*, vol. 2, New York: IEEE, pp. 629–636.
- Brotchie, P. R., and Andersen, R. A., 1991, A body-centered coordinate system in posterior parietal cortex, *Soc. Neurosci. Abst.*, 17:1281.
- Dominey, P. F., and Arbib, M. A., 1992, A cortical-subcortical model for generation of spatially accurate sequential saccades, *Cereb. Cortex*, 2:153–175.
- Droulez, J., and Berthoz, A., 1991, A neural network model of sensorimotor maps with predictive short-term memory properties, *Proc. Natl. Acad. Sci. USA*, 88:9653–9657.
- Goldberg, M. E., Colby, C. L., and Duhamel, J. R., 1990, Representation of visuomotor space in the parietal lobe of the monkey, *Cold Spring Harbor Symp. Quant. Biol.*, 55:729–739. ♦
- Goodman, S. J., and Andersen, R. A., 1989, Microstimulation of a neural-network model for visually guided saccades, *J. Cog. Neurosci.*, 1:317–326.
- Goodman, S. J., and Andersen, R. A., 1990, Algorithm programmed by a neural network model for coordinate transformation, in *Proceedings of the International Joint Conference on Neural Networks, 1990*, vol. 2, New York: IEEE, pp. 381–386.
- Mazzoni, P., Andersen, R. A., and Jordan, M. I., 1991, A more biologically plausible learning rule for neural networks, *Proc. Natl. Acad. Sci. USA*, 88:4433–4437.
- Shibutani, H., Sakata, H., and Hyvärinen, J., 1984, Saccade and blinking evoked by microstimulation of the posterior parietal association cortex of the monkey, *Exp. Brain Res.*, 55:1–8.
- Thier, P., and Andersen, R. A., 1991, Electrical microstimulation delineates 3 distinct eye-movement related areas in the posterior parietal cortex of the rhesus monkey, *Soc. Neurosci. Abst.*, 17:1281.
- Zipser, D., and Andersen, R. A., 1988, A backpropagation programmed network that simulates response properties of a subset of posterior parietal neurons, *Nature*, 331:679–684. ♦

Generalization and Regularization in Nonlinear Learning Systems

Grace Wahba

Introduction

In this article we will describe generalization and regularization from the point of view of multivariate function estimation in a statistical context. Multivariate function estimation is not, in principle, distinguishable from supervised machine learning. However, until fairly recently supervised machine learning and multivariate function estimation had fairly distinct groups of practitioners and little overlap in language, literature, and the kinds of practical problems under study.

In any case, we are given a *training set*, consisting of pairs of input (feature) vectors and associated outputs $\{\mathbf{t}(i), y_i\}$, for n training or example subjects, $i = 1, \dots, n$. From these data, it is desired to construct a map which *generalizes well*, that is, given a new value of \mathbf{t} , the map will provide a reasonable prediction for the unobserved output associated with this \mathbf{t} .

Most applications fall into one of two broad categories, which might be called nonparametric regression and classification. In *nonparametric regression*, y may be (any) real number or a vector of r real numbers. In *classification*, y is usually represented as a q -dimensional vector of zeros and ones, with a single 1 in the k th position if the example (subject) came from category k . In some classification applications, the desired algorithm will, given \mathbf{t} , return a vector of zeros and ones indicating a category assignment (“hard” classification). In other applications, it may be desired to return a q -vector of probabilities (that is, non-negative numbers summing to 1) which represent a forecast of the *probabilities* of an object with predictor vector \mathbf{t} being in each of the q categories (“soft” classification).

In some problems the feature vector \mathbf{t} of dimension d contains zeros and ones (as in a bitmap of handwriting); in other problems, it may contain real numbers representing physical quantities. In this article we will be generally concerned with the latter case, since the ideas of generalization and regularization are easiest to discuss when there is a convenient topology (for example, that determined by distance in Euclidean d -space) so that “closeness” and “smoothness” can be easily defined. *Regularization*, loosely speaking, means that some constraints are applied to the construction of the map with the goal of reducing the generalization error (see also REGULARIZATION THEORY AND LOW-LEVEL VISION). Ideally, these constraints em-

body a priori information concerning the true relationship between input and output; alternatively, various ad hoc constraints have sometimes been shown to work well in practice.

Generalization and Regularization in Nonparametric Regression

Single-Input Spline Smoothing

We will use Figure 1 to illustrate the ideas of generalization and regularization in the simplest possible nonparametric regression setup, that is, $d = 1$, $r = 1$, with $\mathbf{t} = t$ any real number in some interval of the real line. The boxed points (which are identical in each of the three panels of Figure 1) represent $n = 100$ (synthetically generated) input-output pairs $\{t(i), y_i\}$, generated according to the model

$$y_i = f_{TRUE}(t(i)) + \epsilon_i \quad i = 1, \dots, n \quad (1)$$

where $f_{TRUE}(t) = 4.26(e^{-t} - 4e^{-2t} + 3e^{-3t})$, and the ϵ_i came from a pseudo-random number generator for normally distributed random variables with mean 0 and standard deviation 0.2. These figures are from Wahba and Wold (1975). Given these training data $\{t(i), y_i, i = 1, \dots, n\}$, the learning problem is to create a map which, if given a new value of t , will predict the response $y(t)$. In this case, the data are noisy, so that even if the new t coincides with some predictor variable $t(i)$ in the training set, merely predicting y as the response y_i is not likely to be satisfactory. Also, this does not yet provide any ability to make predictions when t does not exactly match any predictor values in the training set. It is desired to generate some sort of curve, which will allow a reasonable prediction of the response for any t within a reasonable vicinity of the set of training predictors $\{t(i)\}$. The dashed line in each panel of Figure 1 is $f_{TRUE}(t)$; the three solid black lines in the three panels of Figure 1 are three solutions to the variational problem: Find f in the (Hilbert) space W_2 of functions with continuous first derivatives and square integrable second derivatives which minimizes

$$\frac{1}{n} \sum_{i=1}^n (y_i - f(t(i)))^2 + \lambda \int (f^{(2)}(u))^2 du \quad (2)$$

for three different values of λ . The parameter λ is known as the