Models of Basal Ganglia and Cerebellum for Sensorimotor Integration and Predictive Control

Jerry Huang, Marwan A. Jabri[$]*, Olivier J.- M. D. Coenen[#], and Terrence J. Sejnowski[+]*

* Computer Engineering Laboratory
School of Electrical and Information Engineering
The University of Sydney
NSW 2006 Australia

[+#]Computational Neurobiology Laboratory
UCSD [+] Biology Dept., [#] Physics Dept.
[+] Howard Hughes Medical Institute
The Salk Institute
10010 North Torrey Pines Rd
La Jolla, CA 92037, USA

[#]Research on the Neuroscience of Autism
Children's Hospital Research Center
San Diego, CA 92123

$ Oregon Graduate Institute of Science & Technology
Electrical and Computer Engineering Department
20000 NW Walker Rd, Beaverton, 97006 OR, USA

Email: marwan@ece.ogi.edu

Abstract

This paper presents a sensorimotor architecture integrating computational models of a cerebellum and a basal ganglia and operating on a microrobot. The computational models enable a microrobot to learn to track a moving object and anticipate future positions using a CCD camera. The architecture features pre-processing modules for coordinate transformation and instantaneous orientation extraction. Learning of motor control is implemented using predictive Hebbian reinforcement-learning algorithm in the basal ganglia model. Learning of sensory predictions makes use of a combination of long-term depression (LTD) and long-term potentiation (LTP) adaptation rules within the cerebellum model. The basal ganglia model uses the visual inputs to develop sensorimotor mapping for motor control, while the cerebellum module uses robot orientation and world-coordinate transformed inputs to predict the location of the moving object in a robot centered coordinate system. We propose several hypotheses about the functional role of cell populations in the cerebellum and argue that mossy fiber projections to the deep cerebellar nucleus (DCN) could play a coordinate transformation role and act as gain fields. We propose that such transformation could be learnt early in the brain development stages and could be guided by the activity of the climbing fibers. Proprioceptor mossy fibers projecting to the DCN and providing robot orientation with respect to a reference system could be involved in this case. Other mossy fibers carrying visual sensory input provide visual patterns to the granule cells. The combined activities of the granule and the Purkinje cells store spatial representations of the target patterns. The combinations of mossy and Purkinje projections to the DCN provide a prediction of the location of the moving target taking into consideration the robot orientation. Results of lesion simulations based on our model show degradations similar to those reported in cerebellar lesion studies on monkeys.

1 INTRODUCTION

Robotics faces many difficult issues in motor control. One is sensorimotor integration, the problem of integrating and organizing a larger number of sensors to extract and process useful information for motor control and other tasks. Another is to give robots the capacity to learn and adapt without providing explicit teaching examples. Reinforcement learning provides a partial solution to this problem. Strategies for reinforcement learning and control in multi-goals environment are still being currently developed.

Since the nervous system has successfully solved these problems, it is beneficial to study biological systems to look for inspiration and solutions. In return, testing present models of brain areas to solve current problems in robotics improves our understanding of these brain areas and the limitations of present models.

The development of biologically based models of sensorimotor integration is receiving increasing attention from the robotics and neuromorphic engineering communities (Jabri, Coenen et al. 1997; Jabri, Huang et al. 1998; Doya 1999; Gross, Heinze et al.

1999). Researchers have investigated sensorimotor integration models suggesting roles for the pre-frontal cortex (PFC), the basal ganglia, the ventral tegmental area (VTA), and the cerebellum. It has been suggested that the pre-frontal cortex participate in working motor memory, the basal ganglia in decision-making and the VTA in reward generation (Barto 1995; Houk, Adams et al. 1995; Montague, Dayan et al. 1995; Berns and Sejnowski 1996; Montague, Dayan et al. 1996; Shultz, Dayan et al. 1997).

Three sensorimotor loops have been hypothesized to exist within the brain. One involves the cortical areas including the dorsolateral prefrontal, primary motor, pre-motor, somatosensory and association cortex, the cerebellum and the thalamus; another involves other cortical areas, including primary, pre- and supplementary motor areas, the basal ganglia and thalamus, and a third is an oculomotor loop that involves the superior colliculus, cortical areas including parietal, visual, auditory and somatosensory cortex and the thalamus.

We explore in this paper a predictive sensorimotor system, which hypothesizes a sensorimotor integration loop that involves the cerebellum and the basal ganglia as depicted in Figure 1. The objectives of the exploration are to investigate: (1) functional roles of the cerebellum system architecture, (2) interaction of the cerebellum with sensorimotor associations in the basal ganglia to perform motor control in a real time environment using a microrobot to track a moving target, and (3) to develop associated models and learning algorithms.

We have previously reported a sensorimotor architecture that uses a combined model of a basal ganglia/VTA and cerebellum to track a moving target (Jabri, Coenen et al. 1997) (Jabri, Huang et al. 1998) . In that work, the cerebellum module used backpropagation algorithm to learn to anticipate target position.

In a recent study by Gross, Heinze, Seiler and Stephan (Gross, Heinze et al. 1999) two models based on the functional roles of the basal ganglia, cerebellum and cerebral cortex were proposed in anticipation-based perception for robotic navigation. The integrated system generated several possible actions to interact with the environment, predicted the sensory consequences of such actions, and then these predictions were evaluated to select the most appropriate action for the situation. The authors used a microrobot controlled by the anticipatory model to navigate in a static environment, and showed an improvement in robot's performance in comparison to a reactive system.

In this paper we propose a new sensorimotor architecture and several hypothesis about the role of various cell populations in the cerebellum. In contrast to previous work (Gross, Heinze et al. 1999), our architecture deals with a dynamic environment with a moving target in which the coordinates transformation of the target to a robot reference frame is performed in parts by the cerebellum.

The sensorimotor architecture studied here combines models of the basal ganglia and the cerebellum. Biologically, the cerebellum projects to the motor cortex (MC) and pre-motor cortex (PMC) through the ventral lateral (VL) thalamus (Figure 1). The basal ganglia receives inputs from many cortical areas including motor and pre-motor cortices (MC, PMC) as well as reciprocal connections with the thalamus. In the current architecture, the connectivity has been simplified and the cerebellum makes direct reciprocal connections with the basal ganglia. This is supported by evidence that the cortex can be "directly" influenced by the output of the cerebellum (Middleton and Strick 1997). The VL thalamus is therefore treated as a relay in the model.

The role of the basal ganglia in the present architecture is to map sensory information to motor actions in order to reach a reinforced goal. The role of the cerebellum is to predict the sensory environment some time in the future given the robot's sensorimotor behavior. Both basal ganglia and cerebellum models learn their tasks in real-time and from an initial random state.

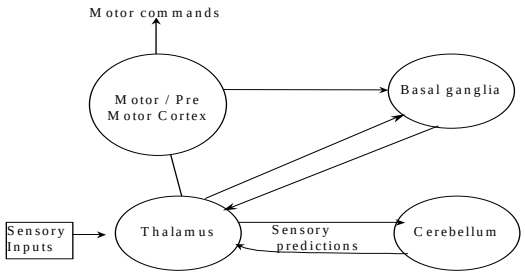


Figure 1. Schematics of the sensorimotor architecture. Sensory inputs are relayed through the thalamus to the basal ganglia and the cerebellum models. The pre-motor cortex, motor cortex and thalamus are modeled as relays, thus direct and reciprocal connections are effectively used between the cerebellum and the basal ganglia.

2 METHODS

2.1 Apparatus

The experimental apparatus consisted of a two-wheeled Khepera microrobot equipped with a top-mounted video camera. The robot was connected by a high-speed serial line to a personal computer (DELL dual Pentium II microprocessor, 400 MHz) where neural computation for learning and motor decisions were made in real time within a multithreated environment written in C++. The robot's environment consisted of an array of 9 infrared light emitting diodes (LEDs) used to simulate a moving target by lighting the LEDs in sequence. Two LEDs of a different color and slightly higher in elevation were placed a small distance from each extremity of the array. These two LEDs were used as landmarks for the robot.

2.2 Visual processing

During the experiments, the robot was placed at a distance from the LEDs in order to have all of the LEDs, including the two landmarks, in the visual field of the camera. The robot attempted to orient itself toward the moving target on the LED array by changing its orientation.The experiments were conducted in the dark to facilitate the detection of the infrared LEDs. Only the luminance component, not the colors, of the images captured by the camera was used. The images were clipped to leave only two horizontal strips that included the target and landmark LEDs. The target LEDs strip was divided into 9 equal regions and the region that contained the most activity was set to ON and all other regions were set to OFF. The activity was determined using thresholding as the intensity captured from the active LED was high and easily recognizable in the dark. The landmark strip was processed for detecting the presence of landmark LEDs, and if present, their horizontal distance from the edge of the visual field were extracted. The activity of the 9 regions of the target strip and the landmark distances were recorded and made available to the rest of the software program. The visual processing as described above was performed in real-time by a separate thread in the software.

2.3 Computational models

The sensorimotor system used in the experiments is shown in Figure 2. An orientation module keeps an estimate of the orientation of the robot using dead reckoning. The current orientation is used to transform the robot-centered target position from the camera visual inputs to a world-coordinate target position independent of robot orientation. The cerebellum module uses the transformed visual inputs and robot orientation to establish predictions of the target location in the visual field at some time in the near future. The sensory predictions produced by the cerebellum are provided as inputs to the basal ganglia model and enhances its abilities in tracking the moving object. The basal ganglia makes motor decisions based on some states of the robot as defined by visual sensory or cerebellar inputs and on the expectation of future rewards on a long-term basis. The ventral tegmental area (VTA) receives an external sensory reward based on visual tracking and computes the predictions of future rewards based on the current state of the system. Hence, whenever the target appeared in the central visual region the reward was set to 1, otherwise it was set to 0. Changes in reward predictions were used to modify the synaptic weights of the basal ganglia network (see below). The basal ganglia network was set up to receive either current visual inputs from the camera or predicted visual inputs from the cerebellum in different experiments. When it received the predicted future visual inputs from the cerebellum, the current visual input from the video camera was not used. The basal ganglia network had two types of output neurons: three motor command neurons and four speed neurons. The three motor command neurons encoded left, right, and stationary motor commands where as the four speed neurons represented different speeds for motor commands ranging from slow to fast.

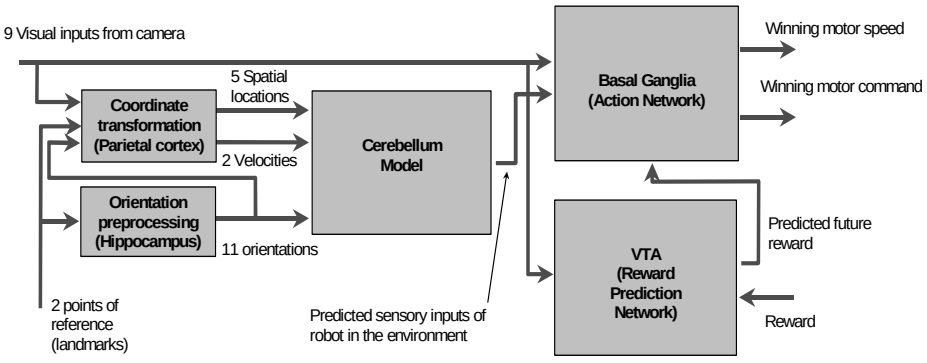


Figure 2. Sensorimotor model used in the experiments. An orientation module keeps an estimate of the orientation of the robot using dead reckoning. The current orientation is used to transform the robot-centered target position from the camera visual inputs to a world-coordinate target position independent of robot orientation. The cerebellum module

uses the transformed visual inputs and robot orientation to establish predictions of the target location in the visual field at some time in the near future. The sensory predictions produced by the cerebellum are provided as inputs to the basal ganglia model and enhances its abilities in tracking the moving object. The basal ganglia makes motor decisions based on some states of the robot as defined by visual sensory or cerebellar inputs and on the expectation of future rewards on a long-term basis. The ventral tegmental area (VTA) receives an external sensory reward based on visual tracking and computes the predictions of future rewards based on the current state of the system. A reward was given by setting its value to 1 whenever the target appeared in the central visual region and was otherwise set to 0. Changes in reward predictions are used to modify the synaptic weights of the basal ganglia network. The basal ganglia network received either current visual inputs from the camera or predicted visual inputs from the cerebellum in different experiments. When it received the predicted future visual inputs from the cerebellum, the current visual input from the video camera was not used. The basal ganglia network has two types of output neurons: three motor command neurons and four speed neurons. The three motor command neurons encode left, right, and stationary motor commands where as the four speed neurons represent different speeds for the motor commands ranging from slow to fast.

2.3.1 Orientation Preprocessing

The orientation of the robot is obtained through dead reckoning. Since the robot execute a motor command every T_{mc} (500 ms), the orientation relative to the start of the experiment *O(t)* at time *t* is calculated using the speed of the motors:

$$O(t) = O(t-1) + T_{mc} * R$$

where R is the angular velocity of the microrobot. The orientation of the robot is encoded using a binary representation over 11 possible orientations as shown in *Table 1*.

1	2	3	4	5	6	7	8	9	10	11
354-5	6-15	16-25	26-35	36-45	45-180	180-315	316-325	326-335	336-345	346-355

Table 1. The table shows the number and range in degrees of the binary units used for encoding the orientation of the robot -- zero is straight ahead. The unit is active if the robot's orientation falls within its range In the model, these units send their output as mossy fibers that project to the cerebellum.

The accuracy of this method was sufficient for the tracking experiments we have conducted. However, it may be improved for longer navigation tasks.

2.3.2 Coordinate Transformation

The coordinate transformation module transforms the robot-centered target position to a world-coordinate target position using a lookup table and a nearest-neighbors technique. The inputs to the module are the current robot orientation and the visual inputs from the camera including the target and the two landmarks. The outputs are represented by a 5-element and a 2-element binary pattern representing, respectively, the position and velocity direction of the target in a world-coordinate system, independent from the actual robot orientation. The transformation was implemented by a lookup table that was constructed using the following procedure. The robot was placed in front of the infrared array representing the moving target and the visual input was recorded for 44 different robot orientations relative to the infrared array. For each orientation, 5 entries in the table were generated; each one was indexed by the landmarks location in the visual inputs and the observed visual pattern. The data associated with each entry is the spatial location of the target in world coordinates. The result is a lookup table with 220 entries of visual–spatial pattern conversion. The lookup table was generated prior to learning in the cerebellum module. The minimum number of binary outputs to represent the possible target spatial locations was 5. More than 5 outputs were tested, but the results showed no difference in performance for the cerebellum module.

2.3.3 The basal ganglia and the ventral tegmental area models

The basal ganglia and ventral tegmental networks are shown in Figure 3. The basal ganglia network is an action network that makes motor decisions based on current visual inputs or predicted sensory states from the cerebellum. The ventral tegmental area (VTA) network is a predictor of expected future rewards.

2.3.3.1 Input and Output Representations

The inputs to both the basal ganglia and VTA networks are the 9 binary visual inputs from the video image. The action network has two different types of output neurons: three motor command neurons and four speed neurons. The three motor command neurons represent left, right, and stationary motor commands. The 4 speed neurons represent different speed of the motor commands ranging in value from 1 to 4 (slow to fast). A winner-take-all mechanism is applied separately to each type of neurons. The winners from each category control the robot together. The predictor network has a single output neuron representing the expected future reward. When the action network receives the predicted future visual inputs from the cerebellum, the current visual input from the video camera is not used (see Figure 2). Learning

The basal ganglia and VTA network models learn sensorimotor associations using predictive Hebbian learning (Montague, Dayan et al. 1995), a form of reinforcement learning. The VTA network output represents a value function similar to the one used in reinforcement learning (Sutton and Barto 1981). The microrobot by exploring its environment gains information on how to change its sensorimotor behavior in order to maximize the cumulative rewards it receives. Since the goal was to pursue the target, the reward was set to1 whenever the target appeared in the central visual region. Otherwise, the reward was set to 0.

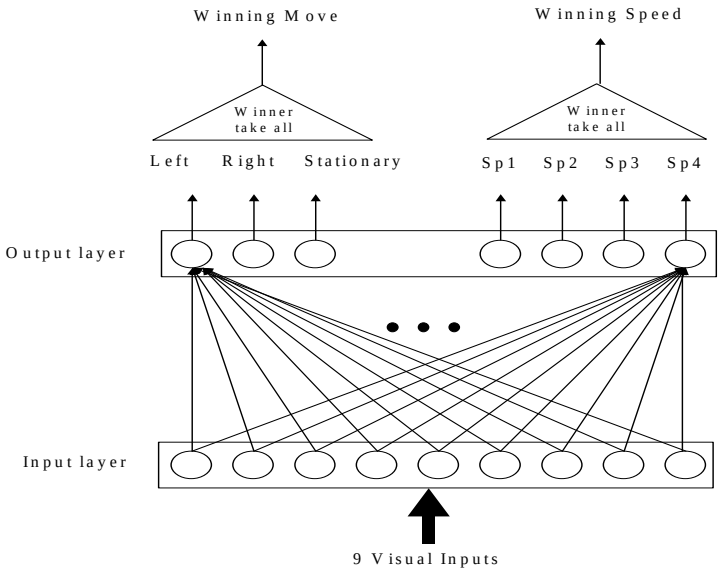


Figure 3. Basal ganglia network. The basal ganglia module attempts to pursue the visual target to obtain a reward based on the current (sensory) state of the robot. The inputs to the network are from the 9 separate regions of the image captured by the CCD video camera. The image is divided into 9 equally spaced binary regions and the region that contains most of the target image is activated. The motor command neurons and motor speed neurons form the output layer. There are 3 motor command neurons: left, right, and stationary and 4 motor speed neurons, Sp1 to 4, ranging from slow to high speeds. A winner-take-all competition is performed on both sets of motor neurons to obtain the winning commands to control the robot. The basal ganglia network models learns to improve the sensory to motor associations to maximize rewards using a predictive Hebbian reinforcement learning algorithm under a definite exploration policy.

The purpose of learning is to maximize the reward received by controlling the motor actions. The VTA network learns by updating a value function *V(t)* that gives the expected value of the cumulative discounted future rewards given the current state of the system. The adaptation takes place using a consistency equation called the temporal difference learning rule (Sutton 1988):

$$\delta(t) = r(t) + \gamma V(t) - V(t-1) \qquad (1)$$

where $\delta(t)$ is the prediction error, *r(t)* is the reward received at time t, *V(t)* is the estimate at time *t* for the sum of discounted future rewards, and γ is the discount factor of future rewards. The prediction error $\delta(t)$ lies between –1 and 1 and is used to improve the value function estimate *V(t)* of cumulative future rewards and to modify the stimulus-response relationship in the action network. The weights in the VTA and basal ganglia networks are adjusted according to the predictive Hebbian learning algorithm (Montague, Dayan et al. 1996) that changes weight values using the correlation between the input and output activity of a synapse (Hebbian) weighted by the prediction error $\delta(t)$:

$$e_{ij} = (1 - \lambda)\, e_{ij} + \lambda\, x_j\, y_i \qquad (2)$$

$$w_{ij} = w_{ij} + \eta\, \delta e_{ij} \qquad (3)$$

where e_{ij} represent the *eligibility* trace for the weight w_{ij} connecting the input *j* to neuron *i*, x_j is the activity of the input *j*, y_i is the output activity of the neuron, λ *is* the discount factor for the trace and η is the learning rate. When λ equal to 1, there is no trace and only the instantaneous input-output correlation is used. The eligibility keeps a memory trace of past correlations and facilitates the correlation with prediction errors to create sequential stimulus-response associations.

Reinforcement learning is closely affected by the exploration of the environment. The exploration policy influences the convergence in finding a solution to the goal of continuously track the object. The policy examines the behavior of the action network and makes a decision on whether to continue with the action generated or to explore further in order to achieve better performance. Our model uses a probabilistic approach to decide whether to explore further:

$$p(t) = \frac{1}{1 + \exp[\delta(t)\,(a\delta(t)^{b} + 1) * 0.5 + V(t) * b - c]} \qquad (4)$$

where *p* is the probability of performing an exploration and *a, b* and *c* are positive constants. The prediction error $\delta(t)$ value ranges from –1 to 1 starting at 0, V(t) ranges from 0 to 1 starting at 0 and the constant were chosen so that a= 20.0, b = 8.0, and c = 4.0.

The above equation attemps to model the exploratory behavior in animals. The constant *a* governs the descente from high to low in an exploration curve. Because $\delta(t)$ starts at zero, without the constant *c*, there is a 50% chance in performing an exploratory move. The constant *c* shifts the curve and makes the model starts with a higher exploration probability. *V(t)* is the expected future reward, that is also the confidence level of the network in performing the movement associated with the stimulus. *V(t)*b* creates a moving window that provides the upper and lower bounds for the curve. As *V(t)* increases, the range of probability shifts to the low end of the curve. So as learning progresses, the probability to perform an exploratory move is reduced and confined to the high region. As the network gets better at performing a task, the probability is shifted to the lower region. The model also has a contingency for the special situations that may occur such as discovering an excellent solution at the beginning of learning or finding out that the learned sequence does not lead to a solution. The term $\delta(t)^{b}$ allows quick transitions to either high or low probability regions to occur by overcoming the confining window.

2.3.4 The Cerebellum Model

Evidence suggests that the cerebellum is involved in timing, motor coordination (Marr 1969) motor learning and sensorimotor integration (Akshoomoff and Courchesne 1992). Cerebellar contributions have been inferred in situations as diverse as timing of the conditioned eyelid response (Perrett and Ruiz 1993), shifting of attention (Akshoomoff and Courchesne 1992), adaptation of the vestibulo-ocular reflex (Akshoomoff and Courchesne 1992) and coordination of eye and hand motor systems (Donkelaar and Lee 1994). Some of these studies also suggest that the cerebellum may be involved in cognitive aspects of information processing (Kim, Ugurbil et al. 1994). Several theories of cerebellar function have been proposed, including the original motor learning theories of Marr (Marr 1969), Albus (Albus 1971) and others. Nevertheless, few of these theories have provided a consistent view of the cerebellum's role in these diverse tasks. The possibility that the common thread across these different tasks is a predictive ability that the cerebellum brings to the central nervous system is one focus of our investigation (Coenen and Sejnowski 1996; Coenen 1998). The predictive functionality of the cerebellum may include prediction of neural signals carrying sensory, motor, and even cognitive information (Akshoomoff, Courchesne et al. 1997).

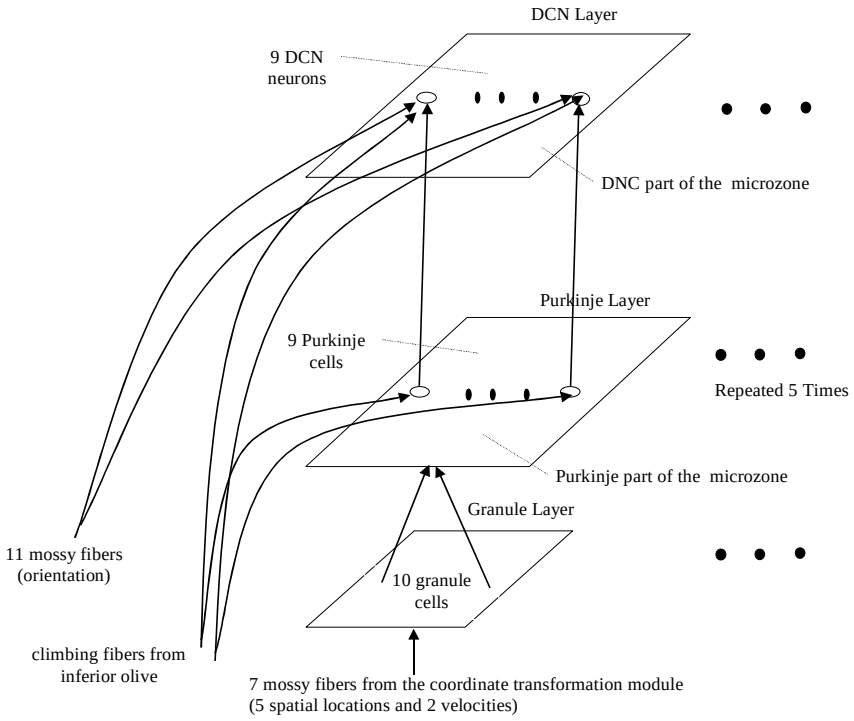


Figure 4. The cerebellum network model. One cerebellar microzone is shown; the structure of the figure is repeated 5 times in the model as indicated. The granular layer receives information from the coordinate transformation module and provides a binary representation for learning in world coordinates that are independent of the robot's orientation.. Two mossy fibers, a spatial location and a velocity fiber, contact a granule cell in a layer that contains 10 granule cells. The Purkinje cells are arranged in a 9x5 matrix, thus defining 5 microzones each having 9 Purkinje cells. Each Purkinje cell in the microzone receives inputs from 10 granule cells and one climbing fiber. Each DCN neuron receives an inhibitory projection from a Purkinje cell and excitatory connections from the 11 mossy fibers of the orientation preprocessing module. The combined outputs of the DCN neurons in a microzone predict the sensory pattern that should appear in the visual field at the next time step. The outputs produced by each microzone are combined to form the predicted sensory pattern, for example, the collective activity from the first neurons in each microzone represents the activity of the first region in the visual field. The climbing fiber from the inferior olive encodes the sensory prediction error, distributes it to the cerebellum and modifies the synapses of neurons that it contacts.

The network model of the cerebellum is shown in Figure 4. The model is based on known biological evidence such as synaptic connectivity, receptor properties, neurotransmitters and synaptic plasticity. The cerebellum model is assigned to predict sensory patterns in tracking the moving target, and assists the basal ganglia model in its sensorimotor action tasks. In the current implementation, the cerebellum model has the ability in the Purkinje layer to predict sensory patterns in world coordinates independent of the robot orientation. The cerebellum also transforms the predicted sensory pattern in world-coordinates back into the robot-centered coordinates of the input visual field in the deep cerebellar nucleus (DCN) layer that receives current robot orientation inputs. The visual field prediction is then projected to the basal ganglia model. This hypothesis suggests that learning to predict the sensory pattern in the cerebellum first occurs in world coordinates and then is transformed to the robot-centered coordinates of the visual inputs. The cerebellum model comprises three interconnected layers: the granular, Purkinje and deep cerebellar nucleus (DCN) layers. The granule layer provides a spatio-temporal basis of mossy fiber inputs for the subsequent layers, the Purkinje layer learns to predict the visual patterns in world coordinates, and the DCN layer combines the robot orientation with the Purkinje layer output to perform the coordinate transformations necessary to predict target location in the robot visual field.

In our framework, the DCN and the cerebellar cortex form together a predictive machine that is under the regulatory control of the inferior olive (IO). The IO provides prediction error signal for learning to the cerebellar cortex and the DCN via climbing fibers. The predictions being made are those of neural activities related to the excitatory inputs reaching the inferior olive. The inhibitory inputs from the DCN to the IO carry a delayed feedback of the predictions that are being established in the cerebellar nuclei (Miall, Weir et al. 1993).

The Purkinje and DCN neurons are organized in microzone (Ito 1984) that consists of a parasagittal strip of Purkinje cells that project to DCN cells that in turn project inhibitory projections to the inferior olive where the climbing fiber, which contacts the same Purkinje and DCN cells in the complex, originates.

2.3.5 Input and Output Representations

There are three mossy fiber inputs to the cerebellum in the model: the target location and velocity direction in world coordinates from the coordinate transformation module and the robot orientation from the orientation module (Figure 2).There are 5 target location mossy fibers, 2 velocity mossy fibers (each representing one target direction of movement: left, and right) and 11 orientation mossy fibers representing the robot's orientation relative to the starting direction. These mossy fibers are assumed to originate from brain structures such as the parietal cortex that is known to encode objects in world coordinates (Andersen 1995) in primates.

The output of the cerebellum in the present implementation predicts the visual activity after preprocessing. The prediction of the sensory pattern is produced by the collective activity of the DCN neurons (see section 2.3.7 below for detail) and represents the future activity of the 9 regions of the visual map at the next time step. This prediction is then sent to the basal ganglia module.

2.3.6 The Granular Layer

The granular layer receives information from the coordinate transformation module and provides a binary representation for learning in world coordinates that are independent of the robot's orientation. Two mossy fibers, a spatial location and a velocity fiber, contact a granule cell in a layer that contains 10 granule cells. The group of 10 granule cells project to each of the five microzone in the network. A granule cell was active only if all its mossy fiber inputs were active, giving it a sparse firing property.

2.3.7 The Purkinje Layer

The Purkinje cells in the model are arranged in 5 microzones, each having 9 Purkinje cells. Each Purkinje cell receives inputs from all granule cells and one climbing fiber (Figure 4). The activity of the Purkinje cell at time t , Pc(t) is given by:

$$Pc(t) = \sum(W_{Gc\text{-}Pc} * Gc(t)) \qquad (6)$$

where $W_{Gc\text{-}Pc}$ are the weights of the granule to Purkinje cell synapses and Gc(t) are the granule cell activities.

2.3.8 Deep Cerebellar Nucleus Layer

The DCN neurons follow the arrangement of the Purkinje cells. There are 5 microzones and each microzone has 9 DCN cells. Each DCN neuron receives an inhibitory projection from a Purkinje cell in the same microzone and excitatory connections from the 11 mossy fibers of the orientation module (Figure 4). The outputs of the DCN neurons in a microzone predict the sensory pattern that appears in a particular region of the visual field at the next time step. The outputs produced by the microzones are combined to form the predicted sensory pattern, for example, the collective activity from neuron 1 in each microzone represents the region 1 in the visual field. The outputs of the microzones are combined using a winner-take-all mechanism. The strongest output in a microzone suppresses other outputs that represents the pattern on the visual field.

The coordinate transformation from world coordinates to robot-centered coordinates in the model is located at the mossy fiber – DCN synapses of the cerebellum. The hypothesis is that this transformation is learned early in brain development, that at that time, the mossy fiber – DCN synapses undergo Hebbian plasticity with respect to climbing fiber activity (equation 7) and that there is no activity from the Purkinje cells during this period, Pc(t)=0. The learning rule for the mossy fiber – DCN synapses is:

$$\Delta W_{Mossy_i\text{-}DCN_j} = \sigma\, Mossy_i(t)\, Cf(t) \qquad (7)$$

where σ is the learning rate (σ=0.001). With this hypothesis, the climbing fiber carries visual input activity to the DCN neuron to learn the transformation. After a developmental period, the climbing fiber changes the signal it carries and encodes sensory prediction error. The prediction error is computed by taking the difference between the one-time-step delayed outputs of the DCN that inhibit the inferior olive and the visual preprocessed inputs that are excitatory inputs to the inferior olive (next section).

The coordinate transformation that we proposed is similar to transformations at the posterior parietal cortex (Andersen 1995). Andersen and colleagues (Andersen 1995) suggested that world coordinates are used in the posterior parietal cortex and that the

transformation is obtained by extracting information from the visual field and eye position. In our model, the visual input activity is predicted from the target location in world coordinates and from the orientation of the robot. The orientation mossy fibers at the DCN synapses modulate the amplitude of the response from the Purkinje cells (in world coordinates) in a manner similar to gain fields. This amplitude modulating gain field implements the coordinate transformation.

2.3.9 Climbing Fiber And Learning Algorithm

After early development (see section 2.3.8), the climbing fibers (Cf) from the inferior olive (IO) encode sensory prediction error. They make extensive contacts with Purkinje cells within a microzone and projects to Golgi cells (not modeled) and DCN neurons. The prediction error is computed by taking the difference between the one-time-step delayed outputs of the DCN that inhibit the inferior olive and the visual preprocessed inputs that are excitatory inputs to the inferior olive. Long-term potentiation (LTP) and depression (LTD) occur in the model at synapses of Purkinje cells and DCN neurons contacted by the climbing fibers. LTD occurs at parallel fiber-Purkinje cell synapses during concurrent activation of parallel and climbing fibers activity when climbing fiber activity is above resting firing rate (set at zero in the model). When the activity of the climbing fiber is below resting firing rate, LTP occurs at these synapses. The weight update rule at parallel fiber-Purkinje cell synapses is:

$$\Delta W_{Gci\text{-}Pcj} = -\alpha\, gc_i(t)\, Cf(t) \qquad (8)$$

and at Purkinje-deep cerebellar nucleus (DCN) synapses, it is:

$$\Delta W_{Pci\text{-}DCNj} = \sigma\, Pc_i(t)\, Cf(t) \qquad (9)$$

where α and σ are learning rates ($\alpha = 0.2$, $\sigma = 0.001$). The learning rate at the Purkinje – DCN synapses is very small in comparison to the parallel fiber – Purkinje synapses. This may reflect the difference in the numerous synaptic contacts that a climbing fiber makes at a Purkinje cell versus the small number it makes at a DCN neuron.

2.4 Experimental Procedures

2.4.1 Intact Experiments

2.4.1.1 Basal Ganglia/VTA Training Procedure

With the basal ganglia/VTA models initialized in random states, the robot was positioned at various angles with respect to a stationary target (LED). Under its exploration policy, the robot explores its environment to learn through reinforcement learning the correct combination of motor command and speed to orient towards the moving target at all times.

2.4.1.2 Cerebellum Module Training Procedure

An infrared array with 5 unique positions is used to simulate a moving target. The cerebellum's task is to predict the movement of the target on the array. The target moves at discrete times through the 5 positions in the array 1 -> 2 -> 3 -> 4 -> 5, back and forth at 500 ms interval between positions.

An experiment with the cerebellum consists of two training phases: (1) Learning the coordinate transformation system in the mossy fiber-DCN synapses, and (2) learning the sensory prediction in the Purkinje cells of the cerebellum.

In the first phase, the robot is placed in front of the 5 target positions, and observes the changes in the sensory inputs for 11 different orientations (see section 2.3.1 and Figure 4) to develop the transformation from world coordinates to perceived robot-centered coordinates of the visual field. This transformation is encoded in the mossy fiber-DCN weights. During this phase, the climbing fibers carry visual input information.

The second phase trains the cerebellum to predict the sensory pattern by placing the robot at an arbitrary initial orientation with respect to the array, but with the array visible for the robot. In this phase, the climbing fibers carry target prediction errors in world coordinates. The weights in the cerebellum model start in a random state. The robot is placed in front of different but fixed target positions, and observes the changes in sensory inputs due to its own movement to develop the transformation from spatial domain (world coordinates) to perceived visual field (robot-centered coordinates). The robot is then placed in front of the array to learn the pattern of the moving target while the robot remains stationary. The cerebellum is trained with the periodic pattern moving back and forth.

During the second phase, sensory predictions were stored in granule-Purkinje and Purkinje-DCN synapses so that both the DCN layer and the Purkinje layer undergo synaptic changes directed by climbing fiber activity. The learning rate at the Purkinje layer was set to be much larger than the learning rate at the DCN layer, making the DCN layer synaptic changes insignificant in comparison to the Purkinje layer changes. This was chosen to preserve the coordinate transformation learned in the first phase.

2.4.1.3 Cerebellum and Basal Ganglia/VTA System

After the sub-systems are trained, the robot is operated to test its abilities in tracking a moving target. The target moves every 500 ms; similarly the robot makes new movements every 500 ms. The cerebellum model performs sensory prediction at discrete time steps, and the basal ganglia module executes movements using the sensory predictions from the cerebellum. When the sensory predictions are used, they form the sole inputs to the basal ganglia.

2.4.2 Lesion Experiments

Many of the brain functional studies are done through lesion experiments, and it is interesting to see how the performance of the system is affected by software lesioning of some of its structures and to compare the results to lesion data on monkeys.

The focus of the lesion studies is on the cerebellum module. The lesion of basal ganglia module will not help in understanding our system behavior, because it simply disables the motor command system. The lesion studies first completely remove the cerebellum sub-system, and observe the behavior of the robot during the same target tracking experiment. Then the performance of the cerebellum model is analyzed by partial lesioning of its internal elements including neurons and synapses.

3 Results And Data Analysis

3.1 Intact Experiments

This section examines the state of the networks and the robot's performance after training for the tracking task. The weight representation of the two modules as well as the robot's behavior are analyzed.

3.1.1 Motor (Basal Ganglia) Sub-System

After training, once a fixed target appears in the robot's visual field, it generates the appropriate motor commands to face the target and then remains stationary. The number of motor commands to reach the target varied with target location. Some locations merely required one motor action to face the target, while other locations required two or more. Nevertheless, the robot ends up facing the target from any orientation. The following 2 figures show the weights of the action networks for the two types of motor neurons. The weights of the motor command neurons are plotted in Figure 5. The weight representation shows that the robot executes a left turn motor command when the visual target appears in the left visual region (position 1, 2 and 3), executes a stationary command when the visual target appears in the central region (position 4, 5 and 6), and executes a right turn for a target to the right (position 7, 8, and 9). Figure 6 shows the weights of the speed motor neurons. The two results show that for example, when the target appears in position 1, the robot turns left at higher speed than for positions closer to the center, because the distance to travel is longer and that all movements last 500 ms.

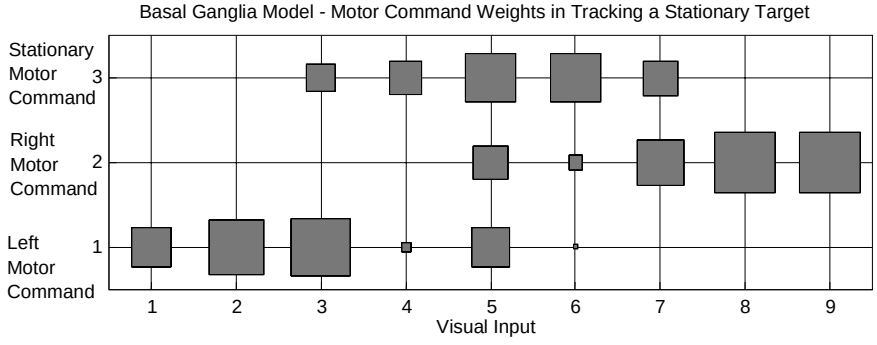


Figure 5. Hinton diagrams for the weights of the motor command neurons of the basal ganglia model in orienting to a stationary target. As indicated by the diagrams, the robot executes left motor command as the visual target appears in

the left visual region (positions 1, 2 and 3), executes stationary command as the visual target appears in the central region (positions 4, 5 and 6), and executes a right turn for a target to the right (positions 7, 8, and 9).

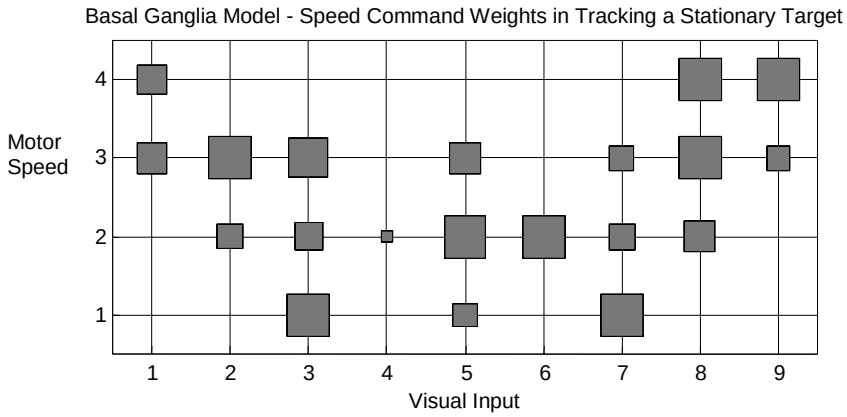


Figure 6. Hinton diagrams for the weights of the speed motor command neurons. The diagrams indicate that the further from the central region the target is, the faster the robot rotates.

Basal Ganglia Analysis

During the training of the basal ganglia module, the robot obtained a reward for reaching the goal state by focusing the target on the center of the input map. The robot displayed one or more movements to reach the goal state at different orientations. This is a consequence of the reinforcement learning algorithm used here. Lets consider two orientations A and B of the robot. Suppose the robot performs a correct move from state A to the goal state, by reinforcement the state A is now associated with certain rewards. Suppose that at state B the robot performs a turn that takes it from an orientation B to A. Because the state A is associated with a certain reward now (V(t) is not zero) and it can lead to the goal state, the movement performed at state B is strengthened (see equation 1-3) with the associated sensory inputs. There is a movement that could transit the robot from state B to the goal state, but the robot usually does not discover it during the exploration stage. It is possible to program the exploration search algorithm such that every possible moves is performed at each state and the best action be selected. However, this would require tremendous amount of processing power and time. The reinforcement learning chooses the best action it discovers in the exploration without having knowledge about an optimal solution. Thus discovering an optimal solution is partly by chance and partly based on the extent of exploration. It is possible to increase the exploration extent of the robot, but it would still depend on the random movements performed during the search. By increasing the exploration time, the chance of discovering an optimal solution increases.

3.1.2 Sensory Prediction (Cerebellum) Sub-System

The first phase of training for the cerebellum module is the acquisition of the coordinates transformation mapping. The second phase of the training is the learning of the predictions.

3.1.2.1 Results for the First Phase of Training

During the coordinate transformation-training phase, the mossy fiber – DCN weights develop to store the robot orientation-to-visual-field coordinates mapping. The weights undergo Hebbian learning while the robot is observing the possible target positions in 11 different orientations. DCN neurons are grouped into 5 microzones as depicted in Figure 4. Each group is responsible for transforming a target location in world coordinates into a robot-centered coordinates of the visual field.

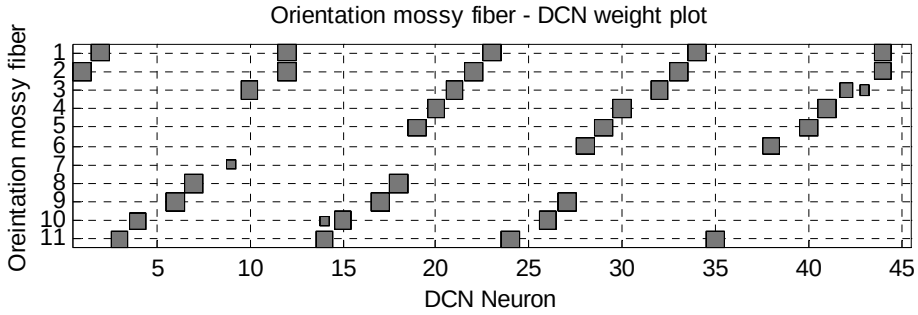


Figure 7. The mossy fiber – DCN weights after coordinates transformation training. The 11 mossy fibers from the orientation module project to the DCN neurons. There are 5 regions each containing 9 DCN neurons cover, for most, the full range of orientations (1 to 11): neurons 1-9, 10-18, 19-27, 28-36, 37-45. Each region is part of a microzone that represents one target location in world coordinates (see Figure 4). A particular orientation mossy fiber is active when the orientation of the robot is within its range (Table 1). For example, when the orientation mossy fiber 1 is active (robot facing straight ahead), this mossy fiber invokes a response in all 5 regions or possible target locations as indicated.

3.1.2.2 Results for the Second Phase of Training

The second phase of training is about learning to predict the moving target. This information is stored at the parallel fiber-Purkinje synapses during LTP-LTD learning. Training was stopped when the mean squared error at the DCN outputs decreased close to zero (0.01 in experiments).

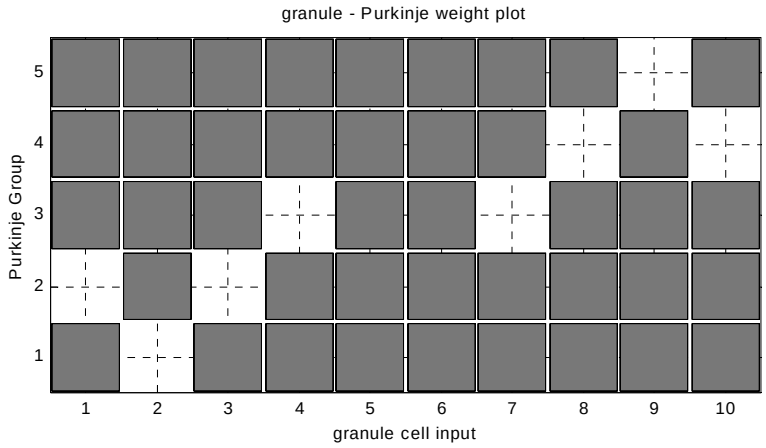


Figure 8. Hinton diagrams of the granule–Purkinje cells synaptic weights. 10 granule cells project to all Purkinje cells. The 9 Purkinje cells in each microzone have identical weights and these are plotted per Purkinje group in the figure. This is a consequence of the fact that the Purkinje cells in a microzone all receive projections from the same granule cells and climbing fiber inputs.

During training, the granule - Purkinje cells weights decreased from the original high values, and the effect is seen at the DCN layer. As the inhibition from Purkinje cells on the DCN neuron decreases; the excitatory inputs from the mossy fiber on the DCN neuron becomes significant, and the output DCN neurons start to encode visual target positions.

The coordinate transformation from world coordinates to robot-centered coordinates is located at the mossy fiber – DCN synapses. The coordinate transformation is ineffective until the Purkinje layer learns the pattern and reduces the corresponding inhibition on the DCN neuron. The gain fields from the mossy fiber - DCN synapses then starts to influence the output of the DCN neuron. At the Purkinje layer, the target is represented in world coordinates (spatial domain) with the 5 microzones. The spatial domain representation is changed into the visual domain representation learned during the first learning phase.

3.1.2.3 Cerebellum Analysis

The proposed functionality of the layers of the cerebellum is verified by examining the activities of the neurons. The granule cells were modeled to represent a binary basis with sparse firing property. The plot in Figure 9 shows the activities of the granule cells in the first group when the moving target just completed 2 cycles. Each cycle will activate 8 different granule cells that were distinguished by their unique spatial location and direction mossy fiber connectivity.

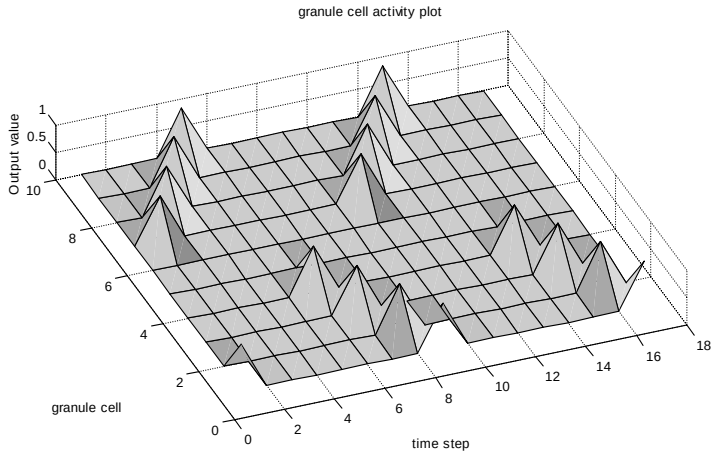


Figure 9. The granule cells are organized in a way such that 1-5 granule cells are tuned to left movement, and 6-10 granule cells are tuned to right movement. Each line represents a different time line, and thus at each time step, there is only one granule cell firing.

The response patterns of the granule cells activate the Purkinje cells that originally (before learning) have large weights. Through the combination of LTP and LTD, the Purkinje cells are tuned to specific granule cells. Figure 10 shows the Purkinje cell response with respect to the granule cells 1 and 2. The response of a Purkinje cell is tuned to reduce its activity for a particular spatial location and target movement, resulting in a reduction of the inhibition on the associated DCN cells.

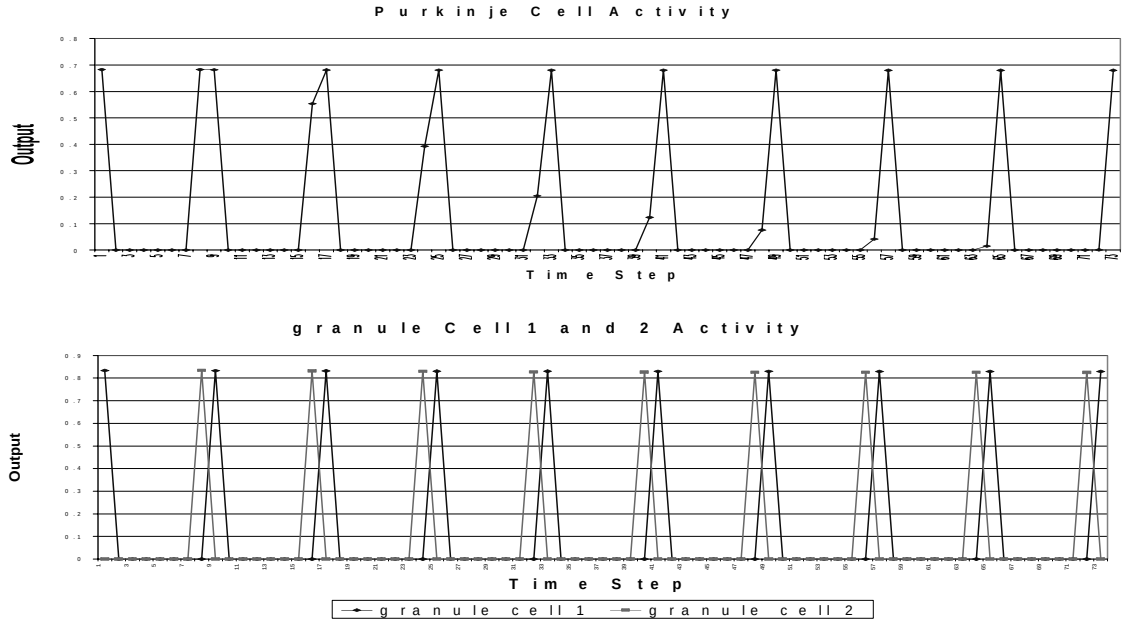


Figure 10. Purkinje cell activity (Up), granule cells 1 and 2 activities (Below). Initially a Purkinje cell responds equally to the two granule cells. After learning, its response is reduced for granule cell 2.

We will use an example to explain the interactions between the layers of the cerebellum. Suppose the spatial location of the target is at the position 2, moving toward position 1, and the orientation of the robot is zero degree with respect to the starting orientation. There visual input will invoke a response in the granule cell 2 in all groups. In Figure 8, the granule - Purkinje weights are unchanged except for microzone 1. The result is due to learning that indicates that the predicted spatial location will be in location 1. The orientation mossy fiber 1 is active to represent zero degree in Figure 7. This mossy fiber will invoke response in all 5 possible target locations as indicated. What we are interested in are the responses in the microzone 1, because all other responses in other spatial locations will be inhibited by the strong activities of the Purkinje cells, expect for the Purkinje cell in microzone 1. According to the recorded data, the orientation mossy fiber activates the DCN neuron 2 in microzone 1, and with decreased response from the corresponding Purkinje cell in microzone 1: the DCN neuron's response indicates the transformed spatial location on the visual map. Thus, the DCN layer changes the spatial location predicted by the Purkinje cells into the visual field domain.

3.1.3 System Performance

The most desirable outcome is for the robot to move in synchrony with the moving target without any delay by keeping the target at the center of visual inputs (positions 4, 5, and 6). The behavior of the robot exhibited during the experiment is expected for a predictive system. The predictive capability allows the robot to show anticipation of future target positions. Thus, the robot is able to move ahead of the moving target. The robot is moving to a large extent in synchrony with the target. This behavior, although much simplified here, is analogical to that observed in monkeys in similar tasks. Monkey shows anticipatory behavior when tracking a moving target using their eyes. The study indicates that the monkey's eye movement is moving at the same speed or faster than the periodic moving target without any delay (Bahill and McDonald 1983).

In terms of target centering performance (see Table 3), the target is appearing over 95% of time at position 3,4,5,6, and 7. Position 3 and 7 are adjacent regions to the central regions. The target appears at the center region over 80% of the time and 15% of time at position 3 and 7. These results give an indication on how often the target is maintained within the goal state. Together with the observation about the movement of the robot relative to the moving target, the robot was displaying an anticipatory behavior while performing the task as the monkey in smooth eye pursuit of a predictable target.

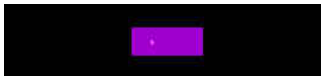
i	ii	iii
		
iv	v	vi
		
vii	viii	ix
		
x	xi	xii
		

Table 3. The above sequences show the robot's view when under the control by the cerebellum - basal ganglia system. The target is appearing over 95% of time at position 3,4,5,6, and 7. Position 3 and 7 are adjacent regions to the central regions. The target appears at the center region over 80% of the time and 15% of time at position 3 and 7. The percentage result only gives an indication on how often the target is within the goal state. As it shows, the target appears at the center of visual field (gray region), expect for frame 5 and 7. As the target is very close to the center of visual map, this error may be due to the generalization of motor commands for different location within a visual section and also due to the motor error associated with hardware.

The performance of the system noticeably degrades with time. The orientation obtained from the on-line odometer calculations deviates from the real robot's orientation. This is due to the error building up from the robot's motors and the calculation of orientation during rotations. The error eventually leads to the target falling outside of the central regions. This only occurs when there is a large discrepancy between the calculation and the actual orientation.

3.2 Lesion Results

3.2.1 Total Cerebellum Sub-System Lesion

The cerebellum module was lesioned from the system, leaving only the trained basal ganglia module to control the robot in the tracking task. With this lesion the robot attempted to tract the target, but always lagged behind it. By the time the robot finished a movement, the target had moved away. The target rarely appeared at the center of the input visual field, and failed to receive a reward most of the time.

Lesion studies on monkeys support the cerebellum's role in providing predictive capacity in movement related tasks. Lu and colleagues (Lu, Hikosaka et al. 1998) performed experiments with Muscimol injections into monkey's dentate nucleus after training them to perform sequential button press tasks and anticipatory saccades. They found an increase in the number of errors while performing the tasks. The anticipating saccadic eye movement performance decreases dramatically. In study by Takagi and colleagues (Takagi, Zee et al. 1998), the dorsal cerebellar vermis was lesioned while sparing the deep cerebellar nuclei. They also observed an increase in the number of errors while performing saccades and an absence of anticipatory behavior. An increase in the number of errors while performing the predictive tasks and a lost of anticipation were observed in these experiments on monkeys and were also observed in the current robot lesion experiment. Table 4 shows the target position during tracking from the robot's perspective. The table shows that movements of the robot after cerebellar lesion are worse than with an intact cerebellum.

I	ii	iii	iv
			
V	vi	vii	viii
			
Ix	X	xi	xi
			

Table 4. Target position during tracking viewed after cerebellar lesion. The frames show that the target rarely appears at the center of visual field (gray region). The basal ganglia module alone does not exhibit the required anticipatory behavior to track the target.

3.2.2 Partial Cerebellum Lesion

The cerebellum network is further examined by selectively removing internal elements in each layer. The effect of the lesion is observed through the prediction errors that the network generates. The cerebellum module was trained to produce close to zero mean squared error in predicting the pattern of the moving target before the lesion. The robot stayed stationary during the lesion analysis. The partial lesion studies include removing of granule cells, Purkinje cells, and mossy fiber – DCN synapses.

3.2.2.1 Granule Cell Lesion

The connection from the granule cells to the first Purkinje cells microzone (9 cells) were severed (referred to as *granule cells lesion* below) so that they did not receive any inputs. Figure 11 shows that the prediction error is high, and that the cerebellum

ceased to function normally. Without the granule inputs, the Purkinje cells no longer exert an effective inhibition on the DCN cells. As a result, the DCN cells are strongly excited by the mossy fibers, and their activity does not decrease. The periodic drop in the error indicates that the cerebellum output matches a visual pattern, in this case, when the expected target is at location 2 in the visual field. Figure 12 shows the DCN output pattern for each time step of the cycle. Notice how the DCN neuron 2 is active at each time step. The error is mainly due to an increase in the activity of the DCN cell 2 response. The rest of the cerebellum still function normally, however the predicted output is a mixture of correct and incorrect patterns.

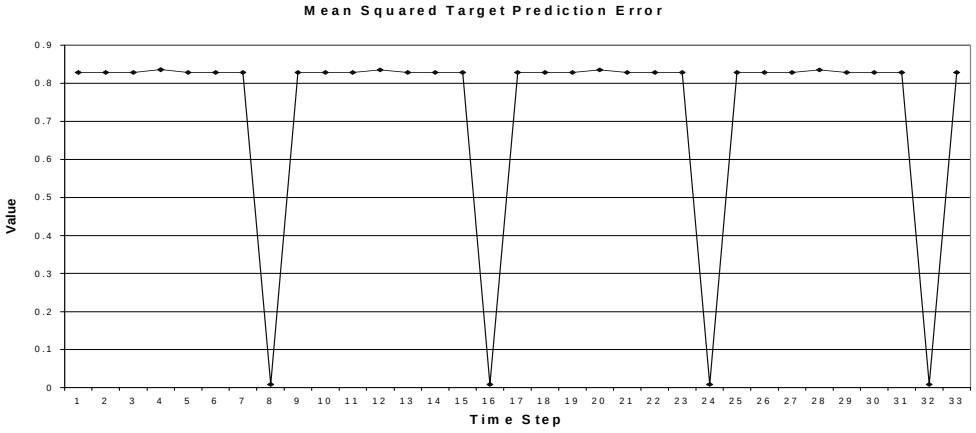


Figure 11. The mean squared target prediction error after granule cells lesion. The errors emerge from a lack of inhibition from the Purkinje cells that are normally activated by the lesioned granule cells. The figure shows that the prediction error is very high, and that the cerebellum ceased to function normally. Without the granule inputs, the Purkinje cells no longer exert the same inhibitory effect on the DCN cells. This results in some DCN cells being strongly excited by the mossy fibers with the activity persisting. The periodic drop in the error indicates that the persisting output of the cerebellum model happens to match the pattern in this case, with the target at visual location 2 (DCN neuron 2, see Figure 12).

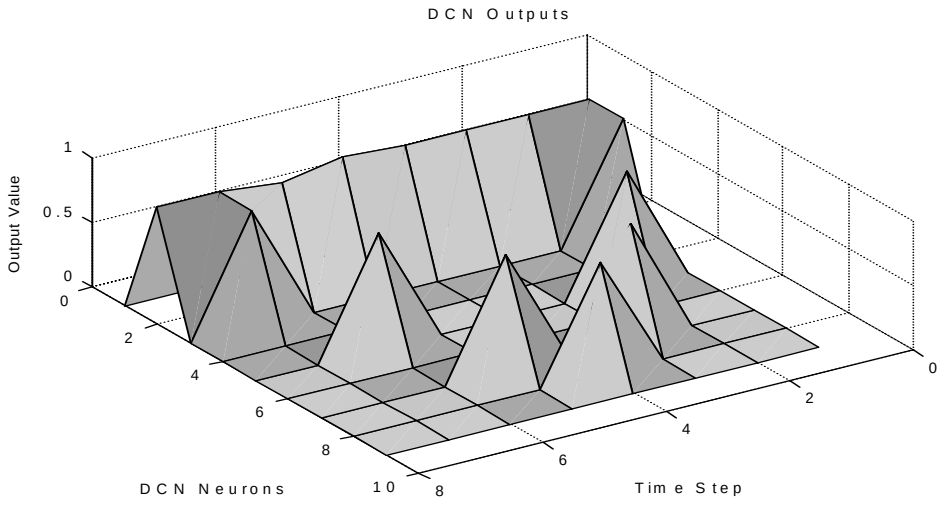


Figure 12. DCN activity after granule cells lesion. The DCN neuron #2 remains active at every time steps causing the constant error seen in Figure 11.

3.2.2.2 Purkinje Cell Lesion

The 9 Purkinje cells in the first microzone are lesioned. These cells receive projections from the previously lesioned granule cells. The result is very similar to the previous section as expected. The granule cells lesion effectively removes subsequent Purkinje cells in the next layer.

3.2.2.3 Mossy-DCN Projection Lesion

The mossy fiber projections of the first 9 DCN neurons corresponding to a target in spatial location 1 were lesioned and the response of the system was observed. The system showed periodic increase in mean squared error that occurred when the expected target position was at spatial location 1(see Figure 13). The lesion of mossy fiber – DCN projections prevents excitatory signals from reaching the DCN neurons, leaving only the inhibition from the Purkinje cells. As a result, the DCN cells are not active, even though the activity of the Purkinje cells is reduced to indicate the next target location. The predicted target location cannot be transformed into visual coordinates without the mossy fiber – DCN coordinate transformation illustrated in a previous section.

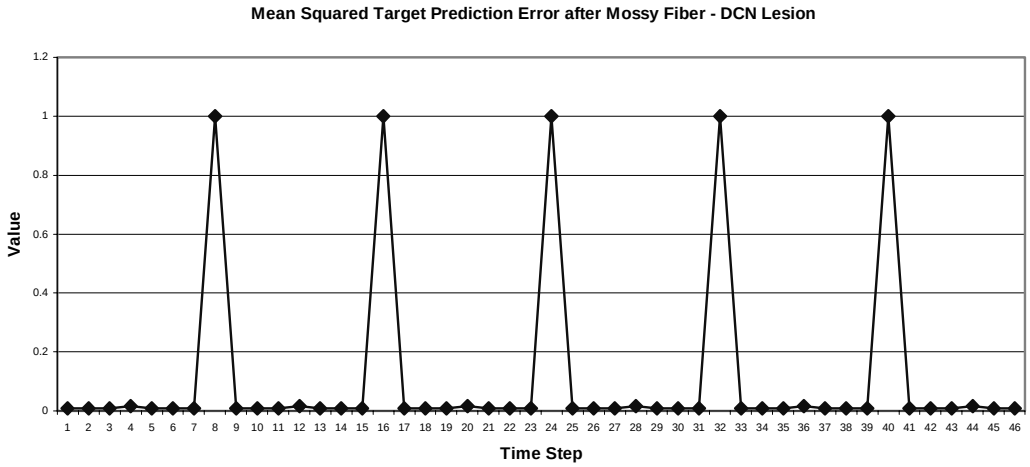


Figure 13. Mean squared target prediction error after mossy fiber – DCN lesion. The system shows periodic increase in mean squared error that occurs when the expected target position is at spatial location 1. The lesion of mossy fiber – DCN projections prevents excitatory signals from reaching the DCN neurons, leaving only inhibition from the Purkinje cells. As a result, the DCN cells are not active, even though the reduced activity of the Purkinje cells indicates the next target location correctly. The next target location in world coordinates cannot be transformed into visual coordinates without the mossy fiber – DCN synapses.

3.2.3 Lesion Analysis

Although all lesions strongly reduced the predictive performance of the cerebellum, the different lesions have different effects on the basal ganglia. In the case of the granule cells and the Purkinje cells lesion, the cerebellum produces both correct and incorrect predictions causing the basal ganglia to receive some erroneous predictive signals from the cerebellum. With the current basal ganglia model, both signals are used to generate motor command in a winner-take-all fashion, but it is unclear how the primate brain could deal to such situations. Lesion of the mossy fiber – DCN projection has a different consequence, it effectively inhibits the cerebellum output stage due to a lack of excitatory drive to the DCN neurons. Partial lesion of the mossy – DCN connection affects certain spatial predictions, leaving some correct and some incorrect. This allows the part of the cerebellum unaffected by the lesion to function normally.

4 DISCUSSION

Our model of the basal ganglia/VTA uses a reinforcement learning algorithm to learn the sensorimotor mapping to track a moving target using the predictive ability of the cerebellum.

In the model, each microzone of the DCN neurons represents the whole visual field and are combined using a winner-take-all mechanism. The specific location in the brain for the integration of the microzones to form a unified representation remains to be specified and investigated. Our model suggests that the integration site is located in a structure the cerebellum projects to; one possible location is the thalamus. In addition to being a relay structure, the thalamus may also combine similar signals projecting to it. The same visual location represented in the different microzones in the DCN could project to the same neurons in the thalamus to form a unified visual field to be projected to the basal ganglia. In some instances, the cerebellum is known to project directly to some brain structures, like the superior colliculus, without first projecting to the thalamus. This suggests that there may be multiple integration sites. However, another possibility is that the diversity of the cerebellum functionality may require different output representations at the DCN layer. The form of the output representation proposed by the current model for sensory prediction for the basal ganglia might not be plausible for use with other brain structures. For example, in a model of superior colliculus and cerebellum interactions, Wang and colleagues (unpublished results) suggest that the DCN neurons use a different representation for the prediction of eye and head positions that projects to the superior colliculus. In light of the multiplicity of cerebellum's functionalities, different information may requires different representation at the DCN layer.

We proposed and analyzed hypotheses on the roles of the granule cells in constructing a binary basis representation, Purkinje cells in world coordinate predictions, and mossy fiber projections to the DCN in a coordinate transformation from world coordinates to robot-centered coordinates. For this latter hypothesis, it was suggested that the information to the inferior olive could change during development and that different brain areas could provide the training signals for the mossy-DCN synapses to develop the coordinate transformation function. This hypothesis remains to be verified with physiological recordings.

A biologically oriented model of the basal ganglia is the next step in achieving a more realistic interaction model with the cerebellum in multi-objective robotic navigation and planning. The cerebellum model itself can be further developed by including the functionality of other neurons such as Golgi cells, basket cells, and stellate cells.

Exploration and exploitation are two important issues associated with reinforcement learning algorithm. Without an adequate exploration policy, learning is usually slow and often leads to incorrect solution. Exploitation policy makes use of the learned solution without exploration for an effective response. These two issues are at each ends of the scale; the goal is to achieve a balance between these two to maximize both efficiency and correctness of solution. The dual control policy utilized in this paper offers one way to achieve the balance between the two, and is the subject of future research.

5 CONCLUSION

The paper described a sensorimotor architecture that incorporates computational models of the cerebellum, basal ganglia/ventral tegmental area, and orientation and coordinate transformation preprocessings. The architecture operates in real-time (learning and operation) on a microrobot to track a moving target. Hypotheses on the role of the cerebellum cortex and deep cerebellar nucleus in predicting sensory inputs and coordinate transformation were also suggested.

Our experiments and results show that the tracking problem could be solved effectively by decomposing the learning problem into different brain structures, and into different layers of a structure like the cerebellum. The basal ganglia and ventral tegmental area models in our sensorimotor architecture were trained using simple predictive Hebbian reinforcement learning rules, and were able to implement sensorimotor behavior that allowed the robot to track a moving target. The combined architecture shows an anticipatory behavior developing allowing the microrobot to anticipate the position of the target.

The results of our experiments also show that the robot exhibits a behavior similar to that observed in animals when a target is predictable. Software lesions of the cerebellum in our experiments have shown degradation in the anticipatory function of the overall system similar to that observed with monkeys.

Current results are encouraging but more simulations of biologically oriented models of the cerebellum are needed to verify various hypotheses regarding the specific function of the different layers in the cerebellum and of other neural structures as well as to examine the network behavior in terms of exploration, exploitation and dual control.

6 REFERENCES

Akshoomoff, N. A. and E. Courchesne (1992). "A new role for the cerebellum in cognitive operations." Behav Neuroscience **106**(5): 731-738.

Akshoomoff, N. A., E. Courchesne, et al. (1997). "Attention coordination and anticipatory control." Int Rev Neurobiology **41**: 575-598.

Albus, J. S. (1971). "A theory of cerebellar function." Math. Biosci **10**: 25-61.

Andersen, R. A. (1995). Coordinate Transformations and motor planning in posterior parietal cortex. The Cognitive Neurosciences. M. S. Gazzaniga**:** 519-532.

Bahill, A. T. and J. D. McDonald (1983). "Smooth pursuit eye movements in response to predictable target motion." Vision research **23**: 1573- 1583.

Barto, A. (1995). Adaptive critics and the basal ganglia. Models of information processing in the basal ganglia. J. L. Davis, J. C. Houk and D. G. Beiser, MIT Press**:** 215-232.

Berns, G. S. and T. J. Sejnowski (1996). How the Basal Ganglia Make Decisions. Berlin Heidelberg, Springer-Verlag.

Coenen, O. J.-M. D. (1998). Modeling the Vestibulo-Ocular Reflex and the Cerebellum: Analytical & Computational Approaches. Physics Department, University of California, San Diego.

Coenen, O. J.-M. D. and T. J. Sejnowski (1996). Learning to make predictions in the cerebellum may explain the anticipatory modulation of the vestibulo-ocular reflex (VOR) gain with vergence. Proc. of the 3rd Joint Symposium on Neural Computation, Institute of Neural Computation, University of California, San Diego, and California Institute of Technology.

Donkelaar, P. v. and R. G. Lee (1994). "Interactions between the eye and hand motor systems: Disruptions due to cerebellar dysfunction." Journal of Neurophysiology **72**: 1674-1685.

Doya, K. (1999). "What are the computations of the cerebellum, the basal ganglia and the cerebral cortex." Neural Networks **12**: 961-974.

Gross, H.-M., A. Heinze, et al. (1999). "A neural network architecture for sensorimotor anticipation." Neural Networks **12**: 1101-1129.

Houk, J. C., J. L. Adams, et al. (1995). A model of how the basal ganglia generate and use neural signals that predict reinforcement. Models of information processing in the basal ganglia. J. L. Davis, J. C. Houk and D. G. Beiser. Cambridge, MIT Press**:** 249-270.

Ito, M. (1984). The cerebellum and neural control. New York, Raven Press.

Jabri, M., O. J.-M. D. Coenen, et al. (1997). Sensorimotor integration and control. Extended Abstracts of the NIPS*97 Workshop: Can Artificial Models Compete to Control Robots?, Denver.

Jabri, M., J. Huang, et al. (1998). Online sensorimotor integration and control. Proc. of the 9th Austr. Conf. on Neural Networks, Brisbane, Australia.

Kim, S. G., A. Ugurbil, et al. (1994). "Activation of a cerebellar output nucleus during cognitive processing." Science **265**: 949-951.

Lu, X., O. Hikosaka, et al. (1998). "Role of Monkey Cerebellar Nuclei in Skill for Sequential Movement." J. Neurophysiology **79**: 2245-2254.

Marr, D. (1969). "A theory of cerebellar cortex." J. Physiol **202**: 437-470.

Miall, R. C., D. J. Weir, et al. (1993). "Is the cerebellum a Smith predictor?" J. of Motor Behavior **25**(3): 203-216.

Middleton, F. A. and P. L. Strick (1997). ". Cerebellar output channels." Int. Rev. Neurobiology **41**: 61-82.

Montague, P. R., P. Dayan, et al. (1995). "Bee foraging in uncertain environments using predictive Hebbian learning." Nature **377**: 725-728.

Montague, P. R., P. Dayan, et al. (1996). "A framework for mesencephalic dopamine systems based on predictive Hebbian learning." J. Neuroscience **16**(5): 1936-1947.

Perrett, S. P. and B. P. Ruiz (1993). "Cerebellar cortex lesions disrupt learning-dependent timing of conditioned eyelid responses." Journal of Neuroscience **13**: 1708-1718.

Shultz, W., P. Dayan, et al. (1997). "A neural substrate of prediction and reward." Science **275**: 1593-1599.

Sutton, R. (1988). "Learning to predict by the methods of temporal differences." Machine Learning **3**: 9-44.

Sutton, R. S. and A. G. Barto (1981). "Towards a modern theory of adaptive networks: Expectation and prediction." Psy. Review **88**(2): 135-170.

Takagi, M., D. S. Zee, et al. (1998). "Effect of lesions of the oculomotor vermis on Eye movements in primates: Sacades." J. Neurophysiology **80**: 1911-1931.