BIOLOGICAL MODEL OF VISION FOR AN ARTIFICIAL SYSTEM THAT LEARNS TO PERCEIVE ITS ENVIRONMENT

Michael R. Blackburn and Hoa G. Nguyen
Undersea A.I. and Robotics, Code 943
Naval Ocean Systems Center
San Diego, CA 92152-5000

Abstract: A computer algorithm is described which implements models of the biological visual mechanisms of the retina, thalamic lateral geniculate and perigeniculate nuclei, and primary visual cortex. Motion and pattern analyses are performed in parallel and interact in the cortex to construct perceptions. We hypothesize that motion reflexes serve as unconditioned pathways for the learning and recall of pattern information. The algorithm demonstrates this conditioning through a learning function approximating heterosynaptic facilitation.

Introduction

Our objective is to design an artificial vision system for use in robotics applications. Because the desired performance is equivalent to that achieved by nature, we anticipate that our objective will be accomplished most efficiently through modeling aspects of the neuroanatomy and neurophysiology of the biological visual system.

Information enters the biological visual system through the retina and is passed to the lateral geniculate and optic tectum. The lateral geniculate nucleus (LGN) also receives information from the cerebral cortex and the result of these two inflows is returned to the cortex. The optic tectum likewise receives the retinal information in a context of other converging signals and organizes motor responses. We do not as yet have a model of the optic tectum, nor do we have any motor capability. Biological systems are of course complete with ocular muscles, neck muscles and other means of orienting to--and acting on--visual stimuli. It is the control of these motor events to which the central nervous system is dedicated. We must keep this in mind and endeavor to provide such capability when designing artificial neural systems, otherwise our products will not achieve independence from our own perceptual motor capabilities.

Physiological and Behavioral Considerations

The visual system is divisible into two major subsystems: pattern processing and motion processing. The processing in both subsystems begins in the retina, remains separable but with one significant cross action in the thalamus, and continues on parallel but interacting networks in the cerebral cortex. A comparison of the characteristics of the pattern and motion analysis subsystems is available in Stone et al. [1]. The pattern analysis/synthesis subsystem is characterized by elements with sustained activities given a constant stimulus, that are contrast dependent, color sensitive, and slow conducting. The system is relatively insensitive to fast changing stimuli. The motion subsystem is characterized by elements with phasic responses to the onset or offset of stimuli, that are generally contrast independent, insensitive to color, and fast conducting. There is little habituation in motion sensitive elements.

It is a fundamental hypothesis (or bias) of the present work that the motion subsystem, by virtue of its reflexive control over behavior, serves as unconditioned stimuli for processing and learning in the pattern subsystem.

Noton and Stark [2] reported that individuals produced very specific and replicable eye movements (saccades) when examining a pattern during initial learning and during recall. Their results supported an interpretation of perception as a serial process; a recognition of the pattern occurring after accumulation of a necessary amount of data. The accumulation being the serial process. The results of Noton and Stark suggest that the detection of a feature leads to eye movements that are designed to detect additional features confirming the structure of the pattern. To accomplish this, the movements need to be associated with the features. In a serial search, the feature is a conditioned stimulus for the generation of an eye movement. The eye movement results in an additional feature being detected, which is then incorporated into the pattern for confirmation.

The microanatomy of the cerebral cortex supports the conditioning of patterns upon basic motion orienting reflexes. We have summarized some of the standard connections of a cortical column in Figure 1 (see [3] or [4] for an introduction to cortical architecture). The cortex is composed of 6 major layers, numbered from the surface toward the center, that differ in the resident cell types and their connections within and without the cortex. We have used three of the several different types of neurons in our model so far, including pyramidal, spiny stellates and smaller non-spiny stellates (see [5] for a more complete description of cell types and locations). The pyramidal cells are the major source of cortical afferents. Spiny stellates in the visual cortex receive input from the thalamic projection and are located mostly in layer 4. Non-spiny stellates are interneurons with generally inhibitory functions that are found in large numbers in layers 2 and 3 and to a lesser degree throughout the cortical layers. These non-spiny stellates can receive both excitatory and inhibitory synapses, acting as cancelable inverters (the capacity for disinhibition).
Figure 1. Simplified schematic of connections within a column of primary visual cortex. Extrinsic inputs to the cortex are listed on the left. Arrows indicate direction of traffic but do not imply the nature of the influence.

The pattern analysis subsystem is attributed to supragranular networks (with cell bodies located in layers 2 and 3). The pyramidal cells from the supragranular layer send axons out of the cortex generally to terminate in other cortical areas, but also distribute collaterals to the infragranular layer 5 on the way out. These collaterals have considerable horizontal spread. The specific terminals of these collaterals is presently unknown, but they may synapse on either the basal dendrites or upon inhibitory interneurons. If the major input is upon the inhibitory interneurons then the supragranular layers have a means to control the reflex output of the Y-subsystem which is located in layer 5B.

Neurons of the supragranular layer habituate readily. This means that, unlike the X-subsystem of the retina and thalamus, the neurons eventually stop responding to a constant stimulus [6]. One consequence of habituation in the supragranular layers is that patterns and components of patterns have finite persistences. The demise of one pattern allows the emergence of another pattern that might have been competitively inhibited by the first.

Layer 4 is prominent in the visual cortex, receiving input from the lateral geniculate. However, much local processing must occur in this layer because only 10% of the synapses are due to the LGN input [7]. The organization of simple features such as line orientation and direction selectivity are possibly performed at this level. Hubel and Wiesel [8] showed an absence of orientation preference in layer 4, while this exists in supragranular and infragranular layers, suggesting that the orientation preference is being organized by layer 4 networks. In our model, direction selectivity is accomplished in the retina, as it is done in many non-primates. However, with encephalization, it is not only conceivable, but possible to transport those mechanisms to layer 4 of the cortex.

The motion subsystem is attributed primarily to infragranular layers on the efferent side of the thalamic projection. The basal dendrites of the output neurons in layer 5B are readily accessed by the Y subsystem collaterals from layer 4B [9], supporting the role of the Y-subsystem as unconditioned stimuli or reflex pathway. The pyramids of the infragranular layers project primarily to subcortical targets, although some large neurons in layer 6 project to cortical locations [10]. The infragranular layers are invaded by the apical dendrites of the Y-subsystem and the efferent elements. Layer 5 dendrites reach layer 1, while layer 6 dendrites are contained in layer 3B, 4C-alpha and 5A. The apical dendrites of the Y-subsystem are thus subject to influences from other cortical areas as well as pattern activity within the local cortical column. Activity should persist in the apical dendrites due to their high resistance and may thus predict the basal activity representing the unconditioned stimuli. Martinotti type cells send axons from layers 4 and 5 to layer 1 [11] providing a means by which motion activity...
could become conditioned to the pattern activity of layers 2 and 3. The Y-subsystem in addition sends axon collaterals up to the supragranular layers. The terminals of these collaterals are also located in cell-rich zones of layers 2 and 3, suggesting synapses upon either the basal dendrites of layer 2 and 3 pyramidal cells or upon stellate inhibitory interneurons. Mountcastle [12] favors the latter possibility. An inhibitory influence could provide feature selectivity by the Y-subsystem. An excitatory influence could give the Y-subsystem unconditioned control over the X-subsystem.

Our hypothesis is similar to one put forward by Braitenberg [13] but differs in the roles of the apical and basal dendritic activity. Braitenberg suggested that activity in the apical dendrites must be preceded by activity in the basal dendrites for learning to take place. The basal dendrite activity would then predict the apical. This prospect is contraindicated by behavioral data on classical conditioning which shows that the conditioned stimulus is more effective if it precedes the unconditioned stimulus [14]. It has also been shown that basal activity is more potent in determining the eventual activity state of the neuron than activity collected at the apical dendrites [6]. This is consistent with the role of the basal activity as the conditioned stimulus. In agreement with Braitenberg, however, we expect that the physical changes supporting learning and long-term memory will occur in the apical dendrites. We differ from Braitenberg in our expectation that the apical activity must precede the basal and persist long enough to become associated with the basal activity.

How do we come to recognize new images? First of all, significant objects are generally moving, they also may be making noises. Both stimuli command attention. By attention, we mean the focusing of the stimulus source. This process of focusing is reflexive and impresses the image on the pattern analysis subsystem. The motion in this case has preceded the pattern, but only a portion of the pattern that is required to form a perception is placed on the retina. For example, if the image is of a face we may focus first on the nose, or mouth (infants are regularly attracted to these features). After a brief pause, another eye movement occurs and the pattern is replaced. The sequence of patterns is determined by the sequence of eye movements. The sequencing of patterns then is accomplished by the conditioned association between pattern and movement. The appearance of a pattern generates a movement that results in the next needed pattern to build the perception.

Noton and Stark [2] showed us that the movements were not random, but were scene dependent within an individual. At least two things could determine the direction of eye movement following the reception of a pattern. First, motion detectors in peripheral vision could reflexively move the fixation point to that location where new features (high contrast areas or areas of high complexity) are likely to be found. This is a consequence of the movement reflex which responds to the net strength of the input. Second, central influences on the reflex, mediated by the conditioned pathways from layers 2 and 3 upon layer 5 pyramidal cells outputting to the superior colliculus, could direct the eye movements to regions of the visual field where critical features are expected. We propose that the first mechanism dominates in early learning and the second mechanism dominates after learning.

Implementation

Retinal Y-subnetwork for Motion Processing

Our implementation of the motion analysis subnetwork of the retina has been described in Blackburn et al. [15], so it will only be briefly reviewed here. Two mechanisms of direction selectivity were studied. One involving feedforward inhibition, the other feedforward facilitation. The mechanism of inhibition was found to be better at defining slow movements while the mechanism of facilitation was found to function more accurately under moderate noise. Since both mechanisms may participate in the biological retina [16], they have been combined for optimal performance in the present implementation. Our model of the retinal Y-subsystem is shown in Figure 2. The output of the Y-subsystem in the current implementation is an analog value representing, on each program execution cycle, the number of pixels that increased or decreased in intensity relative to their neighbors on either direction of the x or y axes. This analog value is normalized by the averaged output of the bipolar layer. On each cycle, the system provides information on the general location and direction of movement in parts of the image falling on sectors of the visual field (analogous to receptive fields of Y-ganglion cells).

It is known that efferent activity to the retina can enhance the central response in the receptive field by means of disinhibition at the amacrine layer upon which centrifugal fibers terminate [17]. The efferent activity may facilitate motion processing in the retina because stimulation of the corticofugal pathway in the absence of retinal stimulation did not result in a ganglion cell response [18]. In our implementation we have used collaterals from the thalamic Y relay elements as the source of retinal reexertion. Because noise increases the overall threshold levels in our model of the retina, anticipation of movement by the thalamus, indicated by the reexertent signals to the retinal amacrine elements, can selectively reduce the conduction thresholds in advance of an expected movement, thus facilitating reexertion.

Retinal X-subnetwork for Pattern Processing

The primary objective for the retinal X-subnetwork (Figure 3) was to provide information on the relative contrast differences across the receptive field. This was achieved by following the prescriptions for surround inhibition that is performed in the outer plexiform layer [19,20]. Surround inhibition was accomplished by slowly adjusting the potential in the horizontal layer to equate with the potentials in the receptor layers. To do this, horizontal activity had to persist over program cycles. Horizontal elements also spread or accepted potentials laterally in an attempt to equalize their distribution—high potentials were reduced and low potentials were increased independent of the equalization of potential between receptor and local horizontal. Two types of bipolars were used. On-center bipolars resulted from receiving
excitatory potentials from receptors and inhibitory potentials from horizontals, while off-center bipolars resulted from inverting that input. While bipolar potentials may be driven negative, only positive potentials are transferred to X-ganglion elements. An amacrine element collected the on-center bipolar output and provided a normalizing bias on bipolar to ganglion transmission, while another amacrine performed similarly for the off-center pathways. An on-center X-ganglion element produced a higher value of output than other X-ganglion elements when the center of its receptive field had a higher degree of brightness compared to its surround than was present at other points on the receptor surface.

Lateral Geniculate Nucleus of the Thalamus

The thalamus is often considered a relay station for information passing from the peripheral receptors to the cerebral cortex. This concept only partially represents the functions of the thalamus. Singer [21] has provided an exceptionally clear account of the functions of the lateral geniculate nucleus and reticular nucleus of the thalamus and the neuroanatomy supporting those functions. Our model, shown in Figure 4, is constructed from that account. The microanatomy of the motion and pattern subnetworks are generally separate in the thalamus. The retinal X-ganglion cells, which produce tonic firing patterns in response to a sustained contrast, project to the parvocellular area of the LGN. The retinal Y-ganglion cells, which produce a phasic response to a sustained contrast, project to the magnocellular area of the LGN.

Figure 2. Motion analysis subnetwork model of the retina. The circuitry shown is repeated over the two dimensions of the receptor surface.

Figure 3. Pattern analysis subnetwork model of the retina. Shown is the i-th column of the subnetwork, which is repeated over the two dimensions of the receptor surface.
Figure 4. Models for the thalamic lateral geniculate (LGN) and perigeniculate (PGN) nuclei. Input arrives from both the retina and cortex to the LGN. The magnocellular subdivision of the LGN processes motion information while maintaining the retinal discrimination of direction of movement and location of movement on the receptor surface. Lateral (mutual) inhibition of relay elements reduces noise and ambiguity in the transmitted signal. An inhibitory surround is accomplished in the parvocellular LGN by the convergence of activity on the inhibitory elements that are paired with the relay elements.

Synaptic events recorded from thalamic relay cells after stimulation of the optic nerve show a powerful excitation of about 1 ms followed by a prolonged inhibition lasting about 100 ms, then followed by enhanced excitability [22]. The analog activity of our model could approximate this behavior by using some mechanisms of recurrent and persistent inhibition.

Singer [21] described two sources of inhibition on the thalamic relay neurons, intrinsic and extrinsic. Intrinsic inhibition is very specific (localized to a particular relay neuron) and is the result of input from either the retina or the cerebral cortex. Mechanisms of intrinsic inhibition include binocular interactions (the input from each eye inhibits the input from the other eye covering the same retinal locations), on-center vs. off-center interactions (mutual inhibitions also are found for the same receptive field centers) and Y-subnet control over the X-subnet (the Y-subnet inhibits the X-subnet covering similar receptive fields). In addition there are inhibitory interactions between neighboring relay cells with similar receptive fields (lateral or surround inhibition).

Perigeniculate Nucleus of the Thalamus

Extrinsic inhibition is generalized to a large number of relay neurons and is the result of influences from the thalamic reticular nuclei, the principal of which in the case of the LGN is the perigeniculate nucleus. The extrinsic inhibition is itself under inhibitory control from the cortex and from the mesencephalic reticular formation [22]. Ahlsen et al. [23] reported that the brain stem reticular formation also inhibited the inhibitory interneurons of the LGN. The perigeniculate nucleus (PGN) receives collateral input from most of the LGN-to-cortex and cortex-to-LGN fibers. This input appears to facilitate the perigeniculate role in the inhibition of the relay neurons. A volley of activity in either direction can activate the PGN, which in turn exerts a lasting inhibition of the LGN blocking further transmission. Reticular activation or additional cortical input can block this PGN suppression of the LGN.

The role of the perigeniculate nucleus is relatively minor in the present model. However, it undoubtedly plays an important role in the control of attention. In keeping with Roenneberg's [24] description of the receptive fields of PGN neurons in the cat as being large without inhibitory surrounds, we have modeled the nucleus by a single processing element with a receptive field encompassing the entire receptor surface (Figure 4). The thalamic reticular nucleus also has a high degree of spontaneous activity [22]. We have incorporated this into our model using a leaky membrane equation that depolarizes the membrane. Asymptotic thresholds periodically reverse the current flow resulting in a sine wave modulation of the membrane potential.

The observation of rhythmic bursting in thalamic relay neurons [25] is attributable to the PGN input. This in turn likely contributes to the generation of the alpha rhythm, which is one of the dominant features of the encephalogram (EEG) in resting man. Mental effort or visual input disrupts the alpha rhythm, desynchronizing the EEG. In our model, a
return to the PGN from pyramidal elements in cortical layer 6C reduces the amplitude of the PGN oscillation, reducing the periodic inhibition of thalamic transmission.

**Thalamocortical Loop**

The thalamic input to the cortex is distributed retinotopically primarily to cortical layer 4 (see Figure 1), but also to layers 3 and 6. Pyramidal cells of layer 6 project back to the LGN also retinotopically. The cortical efferents terminate on both the distal dendrites of relay cells and on the inhibitory interneurons [22]. While the strength of the cortical input may be less than that from the retina, through this pathway the cortex can selectively reinforce the pattern of input it receives and suppress other components. A similar idea has been used in the adaptive resonance model of Grossberg [26] and in the alopec model of Harth and Unnikrishnan [27].

**Primary Visual Cortex**

Figure 5 represents our preliminary model of the primary visual cortex. Input to the model cortex is presently limited to the thalamic projections, only because we do not yet have the models for other areas of the brain that communicate with the visual cortex. Similarly, many output elements send axons nowhere as yet. The aspects of the cortex that have been programmed include the layer 4 network that organizes orientation preference, and the pathways that associate orientation preference elements with the eye movement reflexes. In lieu of a brain stem reticular formation, cortical activity is integrated locally and used as a reference for local thresholds and as the information flow from layer 6C to the perigeniculate inhibitory interneuron.

Conditioned learning is accomplished using a model of heterosynaptic facilitation (see [28] for a biological explanation). Our mechanism of heterosynaptic facilitation is shown in Figure 6. Heterosynaptic facilitation is similar to Hebbian learning, forms of which are widely used in neural network studies, in that the changes in the connection weights between the pre- and post-synaptic elements are contingent upon both the level of activity in the pre- and post-synaptic elements. Generally, learning is enhanced by larger values contributed from the pre-synaptic element, but also by larger values resulting in the post-synaptic element. Here is the opportunity for facilitation of one pre-synaptic element by another. An unconditioned stimulus must have a threshold effect on the post-synaptic element, therefore its co-occurrence with a conditioned stimulus will increase the connection weight of the conditioned stimulus, leading to its conditioning and its ultimate ability to activate independently the post-synaptic element. Disfacilitation, or forgetting, is accomplished by the inverse mechanism. Weights are decreased as a function of the post-synaptic activity in the absence of an appreciable contribution from the pre-synaptic element. Weight gains and losses are bounded by values representing the total dendritic capacity of the post-synaptic element and the transmission and/or growth capacity of the pre-synaptic element.

**Model Performance**

The algorithm has been tested in a computer simulation using slowly moving lines with different orientations. The behavior of the model is very much subject to parameters that define the rates of accumulation and loss of potential in the elements. The X and Y subnetworks of the retina respond with potentials indicative of the line location and direction of movement respectively. This activity is integrated in the LGN network, where the initial transmission of the X-activity is blocked by inhibition from the Y-activity. After 5 time constants (program cycle time) the X-subnetwork in the LGN begins to respond to the retinal input. This activity is sent to the cortex and to the PGN. The cortex responds with line orientation elements and the line template is returned to the LGN to reinforce the transmission. Collaterals from the cortex are also sent into the PGN and inhibition of LGN transmission begins to build. Counteracting this inhibition is the potential from layer 6C of the cortex, which acts as an arousal check on the PGN activity. None-the-less, after about 4 time constants the LGN transmission is blocked by the PGN. No new pattern will appear at the cortex until either the image moves or the PGN transmission is blocked by the PGN.
while an old pattern persists in the cortex orientation sensitive elements, the learning algorithm will strengthen their connections to the movement reflexes.

Under the current asymptotic limits on learning and forgetting, the conditioning of pattern to movement reflexes requires about 500 trials. After which, the presentation of a pattern will result in a movement response that would reposition the fixation point to some other part of the visual field. Since there are only 4 effectors to move the fixation point (although the fixation point can be moved by combinations of pairs of these 4 effectors), many patterns can become conditioned to movement in more than one direction. In this case, lateral inhibition resolves conflicts, but recurrent inhibition allows additional movements in different directions subsequently. The unpredictability of element potentials at the time of the movement decision introduces a degree of uncertainty in the direction of gaze with a given pattern after varied learning experiences. With the addition to the model of complex pattern analyzers that feedback into the system, we expect that search patterns will become more directed and replicable.

The system needs many improvements, among which are means to control the extent of ballistic eye movements, and means to associate patterns so that predictions can be made about the features to be encountered as a consequence of a conditioned saccade. So far we have not taken advantage of any of the circuitry in the more superficial cortical layers, nor considered the potential for processing between cortical columns. If our original assumptions are correct, the solutions to the problems of control and prediction should be found in these additional networks.

References


Supported by the Independent Research (IR) program at the Naval Ocean Systems Center. Opinions expressed in this paper are not to be considered as official or reflecting the views of the Department of the Navy.