Elementary Motion Analysis Using a Retina-Inspired Neural Network

Tyler W. Garaas
tgaraas@cs.umb.edu

Marc Pomplun
marc@cs.umb.edu

Visual Attention Laboratory
Computer Science Department
University of Massachusetts Boston
100 Morrissey Boulevard
Boston, MA 02114, USA

ABSTRACT

The processing that occurs in an animal's retina is much more complex than once believed. In addition to highlighting high-contrast areas, the retina also performs a simple motion analysis of the visual field. Moreover, a full understanding of the neural functioning that occurs within the retina is not likely to take place anytime soon. Nevertheless, our present knowledge of the organization of retinal neurons and their responses to stimulation is enough to model meaningful and effective visual processing systems after. In this paper, following a very general, high-level overview of retinal organization and functioning, we demonstrate a neural system that mimics the simple motion analysis that occurs within the retina. The neural system described is modeled after a subset of the connections that are present in most vertebrate retinas.

Categories and Subject Descriptors

I.2.10 [Artificial Intelligence]: Vision and Scene Understanding - Motion

General Terms

Design, Theory

Keywords

Retina Simulation, Motion Analysis, Biological Vision, Computer Vision

1. INTRODUCTION

Over the past century, vision and neuroscience researchers, starting most notably with Santiago Ramón y Cajal, have demonstrated that the retina is much more than a sensory epicenter for the detection of light. In fact, the retina is home to six classes of neurons with many morphologically distinct cell types within each class [3]. Through the precise coordination of all retinal neurons, the retina provides the first stage of the complex processing that enables the meaningful interactions between all visual animals and the world around them.

Our conscious perception of the visual world is far different from the retinal encoding that is known to take place. In the early 20th

century, Hartline demonstrated that some nerve fibers in the optic nerve responded to light, while others responded to the cessation of light [2]. An important discovery soon followed that the responses of some optic nerves was dependant upon the location of the stimulus [2]. Barlow (1953) demonstrated in the frog retina that a stimulus presented in one location would excite a single ganglion cell, the axons of which form the optic nerve, while a stimulus presented at an adjacent location would inhibit the same cell. This type of functionality, termed lateral inhibition, has the effect of increasing ganglion cell responses in areas of high contrast; ganglion cells that respond in this manner are said to be contrast sensitive. Other ganglion cells are known to respond preferentially to motion in a particular direction, and are referred to as direction-sensitive ganglion cells [4]. Thus, it would appear that simple motion analysis of a visual scene can be computed even before the stimulus signal reaches the primary visual centers of the brain. All things considered, it is clear that the retina is a complex center of neuro-circuitry that is responsible for the initial processing and encoding of visual stimuli into a form suitable for higher levels visual processing.

The aim of this paper is to first introduce a very general organization of retinal wiring, and subsequently demonstrate that the creation of neural models based on such wiring have the power to provide functional output for any number of applications. The benefits of creating computer vision systems based upon our knowledge of biological systems are various and many. Two of the most significant, perhaps, include the fact that Mother Nature has solved many of the most difficult problems associated with computing a meaningful representation of the visual world, as all man-made systems created thus far; and second, our own conscious visual experience provides us with a unique insight into the inner-workings of one of the more complex visual systems that Mother Nature has created thus far.

Following the presentation of the general neural layout of the retina, I will demonstrate that a neural model of the retina involving rod photoreceptors, laterally-inhibited bipolar cells, and direction-sensitive ganglion cells is capable of computing the directional analysis of motion in a visual scene. The model utilizes two distinctive synapses found in the retina to facilitate the response of the simulated direction-sensitive ganglion cells.

2. THE RETINA

During early embryonic development, the retina is formed from a laterally-shifted offshoot of the neural tube [4]. Consequently, the retina is considered a displaced region of the brain. Our detailed

knowledge of the retina is no doubt a direct cause of its accessibility when compared to other regions of the brain. The physical location of the neurons, along with their ability to be easily stimulated in a precise manner, has prompted many researchers to use the retina as a means to study neurophysiology in general. Indeed, much of what we have learned from intracellular recordings has come from cold-blooded animals such as the mudpuppy, turtle, and tiger salamander [2].

2.1 Retinal Organization

By and in large, what we know about the neuronal organization of the retina has come from the methodical use of light microscopy of cells stained using the Golgi method. Cajal was first to demonstrate the effectiveness of this method by describing the retinal cells in a large number of species. From his research, and countless others since, we now know a great deal of the cellular organization of the retina. Although there are some differences in retinal organization between species, all vertebrates have the same, basic five-layered retina.

The retina is a layered, thin film of neurons that lines the back of the eyeball. Contrary to what one might first expect, sensory neurons, or photoreceptors (i.e. rods and cones), are situated on the last layer to come in contact with incoming light. Therefore, light must first pass through all other neuronal layers before detection by the photoreceptors. The vertebrate retina shows a very systematic organization consisting of three nuclear layers (outer nuclear layer, inner nuclear layer, and ganglion cell layer) which contain the cell perikarya and two plexiform layers (outer plexiform layer and inner plexiform layer) where the connections between the nuclear layers are contained (see Figure 1). The organization of the retina has a great deal to do with its functionality, as the sensory signal traverses a highly predictable path from the outermost layer (i.e. photoreceptors) to the innermost layer (i.e. ganglion cells).

2.1.1 Outer Nuclear Layer
The outermost layer of the retina, referred to as the outer nuclear layer, consists of two types of photoreceptors in most vertebrate animals. These two types of photoreceptors are rods and cones. The proximal locations and functionality of each type of photoreceptor varies greatly between species. Human retinas contain a dense grouping of cones situated in the center of the retina which account for our acute central vision. The density of cone cells in the human retina falls off rapidly towards the periphery, and is replaced by a greater density of rod cells. Although it is known that photoreceptors contain connections between adjacent cells, little is known about the particular functioning of these connections.

2.1.2 Outer Plexiform Layer
The outer plexiform layer contains the connections between the photoreceptors in the outer nuclear layer and the horizontal, bipolar, and interplexiform cells in the inner nuclear layer. A single bipolar cell contacts either rods or cones, but not both. Connections between cones and bipolar cells come in two primary forms corresponding to the two types of signal processing performed by bipolar cells. The first type of bipolar cell is the on-center bipolar cell, which is excited by light falling on the photoreceptor cells it contacts directly. Conversely, the off-center bipolar cell is inhibited by light falling on the photoreceptors it contacts directly. Connections between the cone cells and on-center bipolar cells take place at a location on the pedicle of the
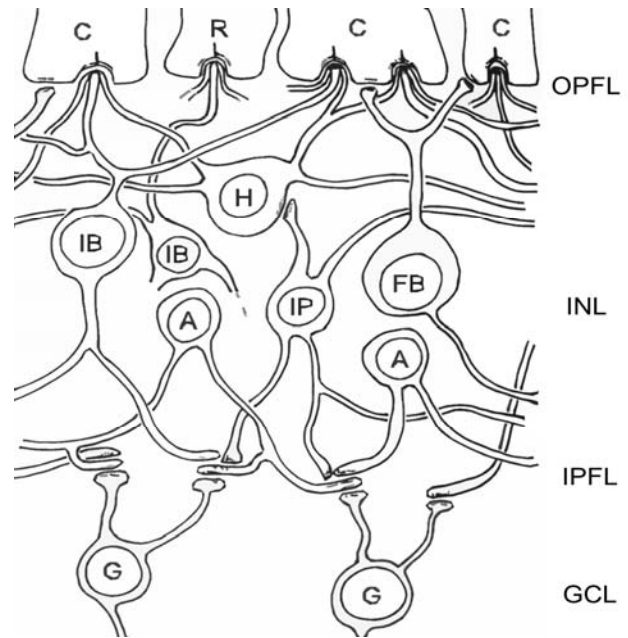

Figure 1. Neural organization in the vertebrate retina. (C) cone, (R) rod, (H) horizontal, (IB) invaginating bipolar, (FB) flat bipolar, (A) amacrine, (G) ganglion, (OPFL) outer plexiform layer, (INL) inner nuclear layer, (IPFL) inner plexiform layer, (GCL) ganglion cell layer

cone cell referred to as a triad, where it is flanked by processes from two horizontal cells. On-center bipolar cells are commonly referred to as invaginating bipolar cells because their dendrites invaginate the cone pedicle. Off-center bipolar cells are connected to the cone pedicle at a basal junction, and are therefore often referred to as flat bipolar cells. Rods are connected solely to on-center bipolar cells and are also usually flanked by two horizontal cell processes.

Aside from being connected to both types of photoreceptors, horizontal cells are also connected to each other through an electrical junction, which effectively increases the area of the visual field on which stimuli can be presented that will affect the response of the cell; this area of the visual field is referred to as a cell's *receptive field*. It is also known that horizontal cells contain a photoreceptor feedback channel, but little is known about the functioning of the channel. Finally, Interplexiform cells, which receive their input in the inner plexiform layer, have axons that are connected to horizontal cells in the outer plexiform layer.

2.1.3 Inner Nuclear Layer
Cell somata of horizontal, bipolar, interplexiform, and amacrine cells lie in the inner nuclear layer of the retina. In some instances, ganglion cells can be located in the inner nuclear layer, and these are referred to as displaced ganglion cells. Morphologically, there are a dizzying array of cell types present in the inner nuclear layer; up to four types of horizontal cells, 11 types of bipolar cells, and 30 types of amacrine cells. Many of these cell types within each class respond to stimulation in similar manners.

2.1.4 Inner Plexiform Layer
The inner plexiform layer contains the connections between bipolar, interplexiform, amacrine, and ganglion cells. Amacrine cells receive their input from bipolar cells and other amacrine

cells, and are presynaptic to interplexiform cells, ganglion cells, and other amacrine cells. Of all retinal cells, amacrine cells exhibit the least consistency in their connections, and their functions are largely unknown. Ganglion cells receive their input from bipolar and amacrine cells, and extend their relatively-elongated axons into the brain via the optic nerve.

In general, bipolar cells terminate on two processes in the inner plexiform layer at a dyad junction. These processes can be two amacrine cell processes, one amacrine cell process and one ganglion cell process, or, though rare, two ganglion cell processes. The organization of cells post-synaptic to bipolar cells appears to have functional significance. Retinas containing a high amount of contrast-sensitive ganglion cells have many more dyad pairings of one amacrine cell and one ganglion cell, whereas retinas containing a high number of ganglion cells with complex responses, such as direction sensitivity, instead have many more dyad pairings containing two amacrine cells [2].

2.1.5 Ganglion Cell Layer
With the exception of a limited number of displaced amacrine cells, ganglion cell somata are the sole inhabitants of the ganglion cell layer.

2.2 Retinal Cell Functions
Research into the activations of retinal neurons has provided researchers with a wealth of information that not only has helped tremendously to decode retinal processing, but also to understand how neurons function in general. The following sections give an overview of what we know of the functions of the different types of neurons in the retina.

2.2.1 Photoreceptors
As mentioned, most vertebrate animals contain two types of photoreceptors, rods and cones. Rods are typically associated with dim-light vision, and cones with bright-light vision. Both types of photoreceptors are hyperpolarized (inhibited) by light and depolarized (excited) by darkness; however, they demonstrate quite different responses to the detection of light. Rods are much more sensitive to light than cones (around 25x in the mudpuppy retina [2]) and depolarize after cessation of light at a slower rate than the cones. Photoreceptors, like most neurons in the retina, generate graded potentials instead of action potentials.

Most vertebrate retinas contain only one type of rod and multiple types of cones. The human retina, like many others, contains three types of cones, which respond maximally to wavelengths of red, green, and blue light and account for our color-vision abilities.

2.2.2 Horizontal Cells
Horizontal cells respond to light with sustained graded potentials, and as with photoreceptors, are hyperpolarized by light. However, unlike photoreceptors, horizontal cells respond to light over a very large area of the visual field, which can be as large as 5 mm or more [2]. This exceptionally large receptive field is undoubtedly a product of the electrical coupling of horizontal cells. In fact, drugs that decrease the electrical coupling of horizontal cells have also been shown to decrease the receptive field of the horizontal cell [4]. It is also believed that the interplexiform cells act on the horizontal cells in a similar fashion [2].
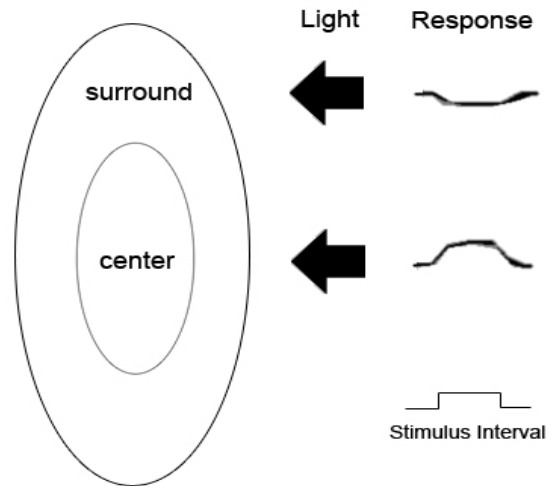

Figure 2. Ideal response of an on-center bipolar cell to light falling in the surround and center of the cell's receptive field. Response line indicates the membrane potential of the cell.

2.2.3 Bipolar Cells
Bipolar cells are the carriers of the stimulus signal from the outer plexiform layer to the inner plexiform layer, and as with horizontal cells and photoreceptors, respond with graded potentials. Bipolar cells respond to centrally-located light in two ways: by depolarizing (on-center bipolar cells) and by hyperpolarizing (off-center bipolar cells). These cells are labeled by their responses to centrally-located light because light that falls on cells adjacent to their centers has the effect of inhibiting their response. For instance, when light falls on photoreceptors that contact an on-center bipolar cell directly (the center of the cell), the cell is depolarized, but if light falls on cells just to the side of these cells (the surround of the cell), the cell is hyperpolarized (see Figure 2). Thus, stimulation of the cell's surround antagonizes the stimulation of the cell's center. Off bipolar cells react in the same manner; that is, stimulation of the center and surround have mutually antagonistic effects on each other. In the case of diffuse light that falls equally on the center and surround of a cell, the response of the center dominates the overall response of the cell, but in an attenuated form.

There is strong evidence to believe that the lateral inhibition of the bipolar cells is achieved through the horizontal cells, as it has been shown that application of a hyperpolarizing current into horizontal cells hyperpolarizes on-center bipolar cells and depolarizes off-center bipolar cells [4]. It is unclear whether the inhibitory effect is achieved by influence on the bipolar cell directly, the photoreceptor directly, or the connection between the photoreceptor and bipolar cell [4]. Interestingly, in some species, the effect of the surround alone is incapable of affecting the bipolar cell directly but requires a center effect to also be present [2]; thus, suggesting either of the latter two methods of bipolar affect.

2.2.4 Interplexiform Cells
As previously noted, it is believed that interplexiform cells are responsible for the mediation of the horizontal cells' receptive field size. It has been suggested that this is accomplished through the use of dopamine to decrease the electrical coupling present

between horizontal cells [4]. Intracellular recordings of the interplexiform cells have shown that the cells fire with sustained, graded potentials.

2.2.5 Amacrine Cells

The responses of amacrine cells generally can be divided into two types, sustained and transient responses, and most retinas will contain many more of the latter than the former. Little is known about the functions of most types of amacrine cells, and in particular, those that generate sustained responses. Amacrine cells that generate transient responses have been shown to be stimulated by the presentation of light, cessation of light, or both; these cells are sometimes referred to as on-, off- and on-off-amacrine cells respectively; such cells always respond to stimulus by depolarizing.

The ability for amacrine cells to transform a sustained, bipolar response into a transient response has been suggested to be mediated by a particular type of synapse found between the two cells. In this synapse, adjacent to the receptor sites of the amacrine cell where input is received, there appears to be a feedback site that could locally inhibit the connection between the two cells. This type of synapse has been termed a reciprocal synapse. Consequently, amacrine cells containing reciprocal synapses would be very sensitive to stimulus changes in their receptive field.

2.2.6 Ganglion Cells

Ganglion cells generate action potentials that travel along their relatively elongated axonal processes which make up the nerve fiber layer and the optic nerve. Ganglion cells come in one of three major types: on-center ganglion cells, off-center ganglion cells, or complex ganglion cells. On- and off-center ganglion cells receive their input from their respective bipolar counterparts and pass on their signal (i.e. on-center bipolar cells depolarize on-center ganglion cells). Diffuse light also evokes an attenuated center-response in on- and off-center ganglion cells.

One of the more understood complex ganglion cells is the direction-sensitive ganglion cell. The direction-sensitive ganglion cell generally responds to motion maximally in one direction, the preferred direction, minimally in the null direction which is usually opposed to the preferred direction, and attenuated in others (see Figure 3). There is much evidence that this response is facilitated by at least a single type of amacrine cell. Another special synaptic connection involving amacrine cells is known as a serial synapse could also play a significant role in the function of direction-sensitive ganglion cells. In a serial synapse, one amacrine cell will synapse onto another amacrine cell which will in turn synapse onto a third cell, which is often a ganglion cell dendrite.

3. Retinal Simulation

Using a small subset of the connections present in the vertebrate retina, it is possible to calculate a form of elementary motion analysis that could be used for a number of ends, such as object tracking, controlling the movement of a camera, or compensating for camera movement.

In the following sections we describe one possibility for using the connections in the vertebrate retina's rod pathway to create a neural model that has the ability to qualitatively describe directional motion within the visual scene of a stationary camera.
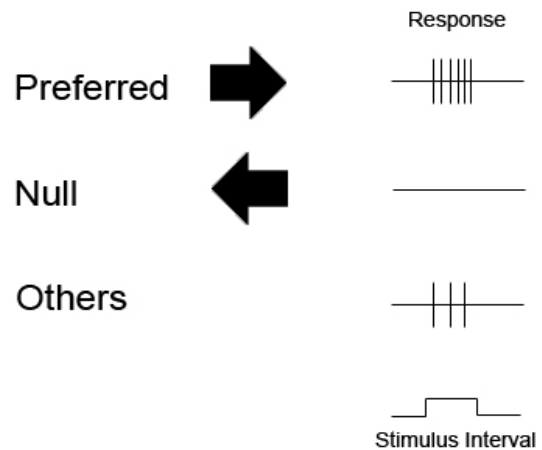

Figure 3. Ideal response pattern of a direction-sensitive ganglion cell with right as the preferred direction. Vertical response lines represent action potentials being generated.

3.1 Neural Network Organization

Input for the neural network was obtained using a Cannon VC-C4 camera at a resolution of 320 x 240 pixels. As is customary with many computer vision applications, grayscale images were used as input. Figure 3a shows a sequence of four images that were used as input for the network. The neural network described follows a single retinal pathway that starts with rod photoreceptors, continues through horizontal, bipolar, and amacrine cells, and finishes with direction-sensitive ganglion cells. In this section, when referring to a simulated cell, it will simply be written 'cell' (e.g. instead of 'simulated horizontal cell', it will be 'horizontal cell'), and when referring to a biological cell, it will be prefixed as such.

3.1.1 Simulated Photoreceptors

Biologically, rod photoreceptors respond to a range of light wavelengths situated between the customary blue and green wavelengths. Somewhat differently, rods in the simulated neural network react to purely intensity values in the image (gray values). The differences here are implemented for convenience only, and could easily be adjusted towards a more biologically accurate model.

Rod photoreceptors make up the first layer of processing in the neural network. Rods are distributed evenly throughout the image on a 1 to 1 basis with pixels in the image, and respond to light (lighter pixels) by hyperpolarizing in the range of 0 to -1, with -1 representing a completely white pixel and 0 representing a completely black pixel. To get an idea of how the rod photoreceptors are responding to images received from the camera, we can simply add 1 to the cell's activation and multiply it by 255 to get a grayscale image showing just that (see Figure 3b).

3.1.2 Simulated Horizontal Cells

Horizontal cells in the neural network are placed every 5 pixels in the input image starting at (5, 5) and ending at (315, 235). Horizontal cells are connected to a neighborhood of rod photoreceptors that extends four pixels in all directions. The size of a cell's receptive field is rather arbitrary in a sense and will require some tuning to reach the most effective value for a

Figure 4. (a) Images captured from a stationary camera and network activations for the (b) simulated rod photoreceptors and (c) simulated horizontal cells.

particular application; this is most certainly the case in nature, as interspecies variability in retinal organization appears to reflect properties fitting their lifestyle [2].

In addition to being connected to photoreceptors, horizontal cells are additionally connected to adjacent horizontal cells in each direction; these connections effectively extend the size of the receptive field, much as they do in biological retinas. The activation of a horizontal cell is computed by equally combining the average of connected rods with the average of connected horizontal cells. The activations of the horizontal cells can be viewed in a manner similar to the photoreceptor cells (see Figure 3c), and have the effect of blurring the input image.

3.1.3 Simulated Bipolar Cells

Bipolar cells are placed every two pixels starting at (2, 2) and ending at (318, 238). Bipolar cells in the neural network, which are always on-center (invaginating) bipolar cells, function much as they do in the biological retina; that is, their center and surround are mutually antagonistic to each other. The center of a bipolar cell consists of connections to photoreceptors one pixel in each direction, and the surround of a bipolar cell is entirely contributed by the single connection to the nearest horizontal cell.

The activation of the bipolar cell is computed as the weighted sum of the center and surround, where the center is the inverted average activation of the connected photoreceptors and the surround is the activation of the connected horizontal cell. The weight of the surround is 80% of the center. This simple antagonism allows the bipolar cell's activation to be much greater in areas of high contrast (i.e. areas where the center is activated with minimal surround activation). As noted, in some species, the surround cannot affect the cell's membrane potential (activation) directly, but instead inhibits the center activation. This could be accounted for by simply limiting the activation of the bipolar cell to positive activations.

To view the activations of the bipolar cell, we simply need to take the absolute value of the activation and multiply it by 255 (see Figure 4a). As shown in Figure 4a, areas of high contrast, such as edges, are activated much more than areas of little contrast.

3.1.4 Simulated Amacrine Cells

Amacrine cells in the neural network are placed directly below each bipolar cell and connect solely to that bipolar cell. However, the synapse made with the bipolar cell is a reciprocal synapse, which effectively computes the amacrine cell's activation. Simply put, the reciprocal synapse is the absolute difference between the current activation of the bipolar cell and its previous activation. This elementary activation function is actually quite powerful, and effectively highlights changes in the visual field. These changes are especially highlighted at areas of high contrast, due to the contrast-sensitive activation function used for the bipolar cell. Amacrine cell activations can be viewed in a manner identical to bipolar cells (see Figure 4b).

3.1.4 Simulated Ganglion Cells

Ganglion cells make up the final layer of processing in the neural network. All ganglion cells in the network are direction-sensitive ganglion cells and only respond to movement in a single direction. Therefore, there are four types of direction-sensitive ganglion cells in the network: left-sensitive, right-sensitive, up-sensitive, and down-sensitive. The different types of ganglion cells differ only by the location of their synaptic connections.

One ganglion cell of each type is placed directly beneath each amacrine cell in the network with the exception of cells on the edge of the network (i.e. right- and left-sensitive cells start under the fourth amacrine cell of each row and finish four cells away

Figure 5. Network activations for the (a) simulated bipolar cells, (b) simulated amacrine cells, and (c) direction-sensitive ganglion cells.

from the far edge). Ganglion cells make synaptic connections to cells in both the null and preferred directions of the cell; these cells can be up to three spots in either direction. Ganglion cells

only make connections using serial synapses. Essentially, the serial synapse combines input from three sources, an amacrine cell in the null direction, an amacrine cell in the preferred direction, and a reciprocal synapse made between the amacrine cell and bipolar cell in the preferred direction. This type of arrangement allows the ganglion cell to receive the activation of the amacrine cell in the preferred direction at two different moments in time.

The activation of the direction-sensitive ganglion cell is computed using Equation 1, where $a1$ is the average previous activation of the connected amacrine cells in the preferred direction, $a1'$ is the average current activation of the amacrine cells in the preferred direction, and $a2$ is the average previous activation of the amacrine cells in the null direction. Consequently, Equation 1 responds to changes that first take place in the null direction and then take place in the preferred. If the change takes place to both cells simultaneously, it will be canceled out by the $a1$.

$$g = a2 * a1' - a1 \qquad \textbf{Equation 1}$$

The activations of the direction-sensitive ganglion cells can be viewed by simply multiplying the activation of the cells by 255 (see Figure 4c).

4. Discussion

The processing that takes place within the retina could be essential for a large amount of our general visual functions. Presently, an estimated 50% of the retinal functions are understood [3], which is perhaps a liberal estimate. At present it is unclear what the full functionality of the retina is and how it is performed, but the retina likely has a hand in processing at least some of the following functions critical for our visual abilities: light-level adaptation, color constancy, and motion perception among others.

The neural network presented, through quite simple, demonstrates the potential of using retina-inspired neural networks to efficiently perform elementary visual processing or form the basis for later, more complex processing. A simulated retina encompassing most of the connections present in the biological retina could form a very powerful start to a larger biologically inspired computer vision system. Furthermore, a computer vision system with a remote camera could effectively reduce the bandwidth required by using local processing similar to that presented in this paper.

5. Acknowledgements

6. References

[1] Barlow, H. (1953). Summation and inhibition in the frog's retina. *Journal of Physiology*, 119, 69-88.

[2] Dowling, J. E. (1987). *The Retina: An Approachable Part of the Brain.* Cambridge, MA, USA. Belknap Press.

[3] Kolb, H (2003). How the Retina Works. *American Scientist*, 91, 28-35.

[4] McIlwain, J. T. (1996). *An Introduction to the Biology of Vision*. Cambridge, UK. Cambridge University Press.