A Hierarchical Artificial Retina Architecture

Alice C. Parker and Adi N. Azar
Electrical Engineering Department, Viterbi School of Engineering
University Of Southern California Los Angeles, CA 90089-2562

ABSTRACT

Connectivity in the human retina is complex. Over one hundred million photoreceptors transduce light into electrical signals. These electrical signals are sent to the ganglion cells through amacrine and bipolar cells. Lateral connections involving horizontal and amacrine cells span throughout the outer plexiform layer and inner plexiform layer respectively. Horizontal cells are important for photoreceptor regulation by depolarizing them after an illumination occurs. Horizontal cells themselves form an electrical network that communicates by gap junctions, and these cells exhibit plasticity (change in behavior and structure) with respect to glycine receptors. The bipolar and amacrine cells transfer electrical signals from photoreceptors to the ganglion cells. Furthermore, amacrine cells are responsible for further processing the retinal image. Finally, the ganglion cells receive electrical signals from the bipolar and amacrine cells and will spike at a faster rate if there is a change in the overall intensity for a group of photoreceptors, sending a signal to the brain.

Dramatic progress is being made with respect to retinal prostheses, raising hope for an entire synthetic retina in the future. We propose a bio-inspired 3D hierarchical pyramidal architecture for a synthetic retina that mimics the overall structure of the human retina. We chose to use a 3D architecture to facilitate connectivity among retinal cells, maintaining a hierarchical structure similar to that of the biological retina. The first layer of the architecture contains electronic circuits that model photoreceptors and horizontal cells. The second layer contains amacrine and bipolar electronic cells, and the third layer contains ganglion cells. Layer I has the highest number of cells, and layer III has the lowest number of cells, resulting in a pyramidal architecture. In our proposed architecture we intend to use photodetectors to transduce light into electrical signals. We propose to employ wireless communication to mimic the gap junction behavior among horizontal cells. These cells could communicate laterally to neighboring horizontal cells through a network of spin wave transmitters and receivers that send magnetic waves over the surface of the first layer of the synthetic retina. We discuss the tradeoffs for having point-to-point connections versus a network on chip in the second layer. We examine the use of 3D CMOS technologies as well as nanotechnologies for the implementation of this retina, considering size, interconnectivity capabilities, and power consumption. Finally, we estimate the volume, delay and power dissipation of our architecture.

Keywords: retina, prosthesis, rod, cone, ganglion cells, horizontal cells, amacrine cells, bipolar cells, architecture

1. INTRODUCTION

Giving sight to the blind is an ambitious goal, but the rewards are obvious. Dramatic progress is being made with respect to retinal prostheses, raising hope for an entire synthetic retina in the near future. Many research groups have proposed various devices. Some of these devices under development, some under clinical trials, and some have been implemented for blind patients.

Now that promising retinal implants are in clinical trials, the possibility of restoring vision is stimulating research involving the development of even more capable neural prostheses in the future. Apart from the extremely difficult process of stimulating the optic nerve with electronics, significant processing goes on in the retina. It would be highly desirable to capture elements of this processing with electronic circuits. Theoretically, if one could reproduce the same output that is created by the human retina using an electronic chip, one could apply it to the optic nerve (ganglion cell axons) and reproduce vision. However, conventional electronics poses several challenges to achieving this goal. In addition to the size and power-consumption issues of conventional electronics, immune system rejection is an issue with prostheses.

In this paper, we propose a bio-inspired 3D hierarchical pyramidal architecture for a future synthetic retina that is designed to fully mimic the overall structure and function of the human retina. We propose an architecture based on the use of several nanotechnologies [1]to facilitate connectivity among retinal cells, maintaining a hierarchical structure similar to that of the biological retina. We divide our system into three layers, and describe each layer's building blocks. We discuss using a choice of nanotechnologies for each layer. We provide loose estimates for volume and power of our proposed system. Although a specific timetable for such a device would be useful, we point to several technological hurdles that must be overcome in order for a fully-functional artificial retina to become a reality, and believe it would be several decades before such a device could be fabricated.

1.1 The Human Retinal Architecture

Connectivity in the human retina is complex, and there is significant processing. This brief and simplified survey is meant to convey the complexity of the anatomy and physiology of the retina, while omitting much of what is known due to lack of space. Over one hundred million photoreceptors (rods and cones) transduce light into electrical signals. Rods are used for low-light vision, and cones are used for daylight. The rods detect light without regard to color, while cones detect red, green, and blue light. These electrical signals are sent to the ganglion cells through the bipolar cells. Lateral connections involving horizontal and amacrine cells span throughout the outer plexiform layer (OPL) and inner plexiform layers (IPL) respectively. Horizontal cells are important for photoreceptor regulation. Horizontal cells themselves form an electrical network that communicates by gap junctions, and these cells exhibit plasticity (change in behavior or structure) with respect to glycine receptors. Horizontal cells communicate with bipolar cells indirectly via feedback to the cone photoreceptors. Bipolar and amacrine cells process the photoreceptors' outputs and forward information to the ganglion cells. On the average, one hundred photoreceptor outputs contribute inputs to each ganglion cell, and multiple ganglion cells respond differently to the same photoreceptors. The ganglion cell processes (axons) form the optic nerve.

The retina performs image processing on the visual scene to extract information. One may look at one aspect of this processing as a step to reduce the ‘bandwidth’ of information flowing to the brain. Amacrine cells are responsible for 70% of the ganglion cells' input, and bipolar cells are responsible for the remaining 30%. Finally, the ganglion cells will spike more frequently if there is a change in the overall intensity for a group of photoreceptors, sending a signal to the visual cortex through the optic nerve. In the human retina, there are about 1.2 million ganglion cells. The connections between the neurons and interneurons in the retina form a complex network with significant differences in connectivity architecture between cells, depending on the cell types.

Cone cells synapse onto bipolar cells. In the dark, cones are depolarized, and release glutamate that inhibits on-center bipolar cells and excites off-center bipolar cells. The release of glutamate causes the on-center bipolar cells to be hyperpolarized. Light causes the cones to hyperpolarize. When the cones hyperpolarize they don't release as much glutamate, and the on-center bipolar cells depolarize. Conversely cones cause off-center cells to depolarize in the dark, and hyperpolarize in light.

Horizontal cells do not synapse onto bipolar cells. They synapse onto cones in the center of the receptive field, depolarizing the cones if the horizontal cells are stimulated by light in the surround. Thus if there is light in the surround, the central cones are depolarized and they release glutamate. In the dark, horizontal cells slightly inhibit postsynaptic cones. In the light, the presynaptic cones are hyperpolarized, causing the horizontal cells to be hyperpolarized, and the postsynaptic cones are depolarized. The horizontal cells and photoreceptors form a complex feedback control network that perform contrast enhancement and allows vision over a wide range of illumination.

Horizontal cells form an electrical network that varies in electrical potential from point to point. The value of the electrical potential at any point depends on the nearby photoreceptors. Whenever a photoreceptor hyperpolarizes, it hyperpolarizes the nearby horizontal cells that in turn depolarize other photoreceptors by lateral inhibition. Depolarizing the photoreceptors is responsible for light adaptation, which is the adjustment of the eye to the degree of illumination. These hyperpolarized horizontal cells may have higher ion concentration than other horizontal cells, leading ions to move from these cells to other cells with lower ion concentration. This specialized intercellular connection is called a

[1] Many applications of nanotechnology to medicine and in particular to neuroscience are being discussed in the literature [23].

gap junction. Gap junctions allows horizontal cells to do spatial averaging of signals on the entire horizontal cell network that will contribute to contrast-sensitive vision over a wide range of illumination.

Each ganglion cell covers a receptive field of photoreceptors. Most ganglion cells are either on-center or off-center. The off-center cells respond with a burst of spikes when the center is turned off. On- and off-center cells are equal in number. Every photoreceptor sends outputs to both types of ganglion cells. Receptive fields of ganglion cells vary from a few minutes (60 min equals a degree of visual angle) in the fovea to the periphery where they cover several degrees (each degree is about .25mm). Within the categories of on-center and off-center ganglion cells there are cells covering large receptive fields (M cells) and small ones (P cells). There are other ganglion cells that don't fall into these categories.

Many of the neurons in the retina, photoreceptors, horizontal and bipolar cells communicate through passive gap junctions rather than through spikes that communicate to post-synaptic neurons through synapses. Photoreceptors produce graded potentials, not spikes, communicating to other cells through gap junctions. Since the gap junctions are essentially analog connections, the effects of neurons communicating with other neurons using gap junctions are subtle and somewhat diffuse.

Many types of amacrine cells and all ganglion cells fire action potentials. Ganglion cells spontaneously produce spikes. Spiking appears tonic and increases with appropriate stimuli. Frequency of firing is related to magnitude of stimulation. The firing rate of a ganglion cell provides a measure of the differences in intensity in the center and surround. The ganglion cells respond to changes in intensity. Increases in center intensity increase firing in on-center ganglion cells. Increases in off-center ganglion cells signals rapid decreases in light intensity in the center.

1.2 Retinal Disorders

Age-Related Macular Degeneration (AMD), and Retinitis Pigmentosa (RP) are two of the most common retinal disorders that may lead to complete loss of vision. According to the Macula Vision Research Foundation (MVRF), AMD is the leading cause of severe and irreversible loss of vision in people over 50 years. AMD can cause irreversible damage to the macula, which is a vital area in the retina, eventually leading to the complete loss of vision. The damage can be due to bleeding, or fluid leakage in the macula. AMD patients see completely dark spots in their vision. Once AMD patients start noticing these dark spots, AMD has already started the irreversible damage to the macula. According to the MVRF, there are 15 million people are suffering from AMD. RP is a genetic disorder that will cause retinal damage, and gradual vision loss. Basically, the disorder damages the retinal photoreceptors that transduce the light into electrical signals. RP symptoms start with night blindness that generally precedes tunnel vision (seeing only certain spots in the visual scene) by many years. It may take many decades for RP patients to become absolutely blind.

There is currently no medical treatment that can completely cure RP. However, taking certain vitamins can reduce the progression of the disease. AMD also does not have an effective cure, but there are new drugs that can be used to cause regression of the abnormal blood vessel behavior that helps to cure the AMD.

As it seems very difficult to cure these two retinal disorders medically within the next few years, a retinal prosthesis is considered to be an alternative solution to restore vision to those who are suffering from irreversible vision loss. Many scientists have investigated various prosthetic devices to restore visions for patients who are suffering from retinal disorders (RP, and AMD) or general visions disorders (optic nerve disorders or completely damaged retina, for example).

The rest of the paper is organized as follows. The first section discusses related work. The next section discusses the challenges of creating a retinal prosthesis implant. Then, we describe our architecture, and discuss the building blocks for various layers.

2. RELATED WORK

In the past eight decades, many scientists have stimulated the brain of various blind patients, and they were successful in producing visual stimuli that allowed those patients to see light or detect the direction of movement of objects. These experiments involved electrical systems that produce electrical signals, and these signals were sent to the brain via different paths including the retina or the cortex. In the first cortical vision stimulation experiment in 1929, Foerster found that when a point at the extreme occipital pole was stimulated by an electrical current, his patient saw a small spot

of light directly. When a point on the medial surface of the left hemisphere just above the calcarine fissure was stimulated, the patient saw a spot of light that moved a little, but was always in the lower right part of the field. In 1988, Dr. Mark Humayun demonstrated that a blind person could see light by stimulating the nerve ganglia behind the retina using an electrical current. This experiment proved that the nerves behind the retina are still functioning well in many cases although the retina has itself degenerated.

Both of the previously mentioned experiments could stimulate the brain to restore a sense of vision. However, they differed in the way the electrical current was applied (cortex or retina), which means that the systems that generate these electrical signals vary in dramatic ways. Implementing a chip that stimulates the cortex directly may help blind people suffering from retinal and non-retinal disorders, which is a big advantage over retinal prosthesis. However, implementing a retinal prosthesis chip seems more feasible in the near term as we need to reproduce the signals created by certain cells in the retina, rather than the more complex task of stimulating the cortex.

Scientists have been working on prosthetic devices that stimulate the human brain to restore a sense of vision for blind patients for many decades [10]. We can classify all the work that has been done in the field of artificial vision into two major approaches: extraocular and intraocular. Extraocular approaches usually involve systems that do not have any implants inside the eye. Cortical prosthesis devices are an example of extraocular approaches. Intraocular approaches include systems fully or partially implanted inside the eye. Intraocular approaches differ in being either epiretinal or subretinal [1]. In epiretinal designs, the prosthesis device will have parts inside and outside the eye. However, in subretinal designs, the device will be fully implanted inside the eye.

Brindley *et al*. developed one of the earliest visual prosthetic devices. Their system was an extraocular cortical prosthesis device [2]. They stimulated the visual cortex for a 52-year-old blind patient by using an array of radio receivers connected to electrodes that are connected to the occipital pole (Posterior part of the cerebral hemispheres). By giving appropriate radio signals, a sensation of light is produced. The Brindley work proved that it is possible to restore vision sensation to patients with a damaged visual pathway. The device required high electrical current, which caused deep pain in the patient's head.

The intracortical approach was proposed to be a better solution compared to the surface cortical prosthesis. As intracortical prosthetic devices have smaller electrodes, they require less electrical current that may reduce the induced pain. The Utah Electrode Array is one of the main intracortical prosthesis projects [9].

Optobionics Artificial Silicon Retina (ASR) is an example of a subretinal intraocular device [3]. Drs. Alan and Vincent Chow of the Optobionics Company developed the ASR chip. It is 2-mm-diameter and 25 μm thick silicon-based subretinal microchip that has 5000 electrodes. Each electrode represents a ‘pixel’ of the visual scene image. The chip uses the incident light to power itself. This passive device is composed of tiny solar cells that, when activated by light, will send an electrical signal through the secondary neurons of the retina down the optic nerve to the brain. A potential problem with such passive devices is that they may generate too little power to be effective. This problem is very prominent in dark areas as there is not enough light to power the chip. In [3] the ASR team reported that the chip was implanted in 6 patients. During a 6 to 18 month follow up period with all the patients, the ASR chips functioned electrically well. Visual improvements were reported, and no signs of implant rejection or infection showed up. The team suggested having larger clinical trial to further evaluate their chip safety and efficiency.

One of the most successful and inspiring retinal prosthesis projects is the U.S. Department of Energy’s (DOE) Artificial Retina Project led by Dr. Mark Humayun of the Doheny Eye Institute at USC. It is a multi-institutional effort to develop an implantable microelectronic epiretinal device that restores vision to people blinded by retinal diseases. The DOE team is working on three models of this chip. These models primarily differ in the number of electrical electrodes that the model has. The goal is to have a large number of electrodes since this will increase the visual scene resolution.

Model 1 of the chip has a 16-electrode array that allows the implanted electronic chip to wirelessly communicate with a camera mounted on a pair of glasses. Model 1 of their artificial retina was implanted in 6 patients whose ages ranged from 56 to 77 at the time of implant between 2002 and 2004. This implant now enables patients to detect when lights are on or off, describe an object’s motion (right, left, up, and down), and count individual items. Patients have very limited vision with model 1 because the implant has only 16 electrodes. Model 2 has 60 electrodes, and it is underway in clinical trials; having more electrodes will help patients have better vision. Model 3 is currently under development. It will have

about 200 electrodes. In [10], the DOE team stated that having at least 1000 electrodes in the implanted chip would allow blind patients to be able to perform quite advanced visual functions like facial recognition and reading.

The DOE artificial retina consists of extraocular and intraocular systems. The extraocular systems consist of a small camera that is built into a pair of glasses, battery unit, and a visual processing unit. The intraocular unit contains the array of electrodes. The extraocular camera captures the visual scene, and transmits it to the visual processing unit that processes the image. The image is transmitted to the intraocular unit using a telemetry system. The intraocular system is a stimulator that stimulates the remaining functional part of the retina.

There are other epiretinal ongoing projects that have achieved great progress including the Boston Retinal Implant project [11, 12, 13] and the Learning Retinal Implant [14]. The Boston Retinal Group is a very promising retinal research group led by Dr. Joseph Rizzo and Professor John Wyatt. The group proposed a subretinal prosthesis. The group fabricated a chip and implanted it on five blind patients. The group is developing a newer chip that is totally intraocular. The Learning Retinal implant is another epiretinal implant developed by Intelligent Medical Implants (headquartered in Zug, Switzerland). The system has both extraocular and intraocular devices. The chip has been under development since 1995. The chip aims to reproduce the receptive field of the ganglion cells using an intraocular chip.

Misha Mahwald *et al*. [7] proposed a visual system that consists of a silicon retina, stereo-correspondence chip and silicon optic nerve. The silicon Retina is responsible for light transduction, and signal processing. The stereo-correspondence chip determines the location of the object using stereovision algorithms. The silicon optic nerve performs digital communication using a technique called address-event representation. The proposed silicon retina had a spatial averaging circuit. It is basically a resistive network, in which each node is connected to six other nodes. Each node potential is going to be affected by the neighboring nodes. This resistive network emulates the behavior of horizontal cells electrical network in the human retina. In the human retina, the horizontal cell layer performs spatial averaging through gap junctions. Furthermore, the retina included an inhibition feedback from the resistive network to the photoreceptors (PNP transistors) using a differential amplifier. This is basically to emulate the adaptation phenomena in the human retina.

Boahen *et al*. proposed a contrast-sensitive silicon retina [6]. They used current-mode CMOS circuits to model the interaction between horizontal cells and cones to mimic the adaptation process that happens in the human eye. Their chip has 90 X 92 pixels. The circuit uses MOS transistors operating in the subthreshold region to emulate gap junctions between horizontal cells and to represent the reciprocal synapses between the horizontal cells and the cones. The design uses PNP transistors as photodetectors. Furthermore, Boahen *et al*. proposed more advanced retinal architectures [8], and Boahen continues to work in this field.

In early writings over a decade ago, Hans Moravec made some predictions concerning the feasibility of an artificial brain and also included the retina in his predictions [20, 21].

3. FUTURE RETINAL PROSTHESIS CONSIDERATIONS

The various decisions that are made by the system architects will affect the size, power consumption, speed of processing and ultimate success of the device. Deciding whether the system is going to be epiretinal or subretinal will make many aspects of the system very different. For example, a subretinal implanted device would most likely be larger than the implanted portion of an epiretinal device.

Regardless of the architecture, these are the general considerations or challenges that will face the architect of a future retinal prosthesis:

1) The artificial retina must deliver visual information in real time, just like the human retina. Having extra delay would mean that the patient would see the scene later than a person with a normal vision. This could lead the patient to make wrong and dangerous decisions like crossing a street at the wrong time.
2) Artificial retina power must not burn or injure the remaining part of the retina or the eye. The human retina consumes a tenth of a watt [19]. An artificial retina should consume no more than that.
3) The artificial retina must be small just like human retina, that is, half a millimeter thick and less than a cm^2 in area.
4) The artificial retina must be biocompatible.

5) The power supply design is challenging. The implanted artificial retina needs to power itself in a way that would not require surgery when we need to replace this power supply. For subretinal devices this is a big challenge since the entire device is located inside the eye, which means more power is required for the implanted device. For epiretinal retinal prosthesis, the problem is relatively easier as we have big parts of the system outside the eye, and we can power those parts using a battery. The DOE team proposed using power telemetry to power the inner part of the retina. For subretinal devices, it is a more challenging problem as the entire device is inside the eye. The ASR team created their chip so it powers itself from incident light.
6) The human retina contains over 100 million photoreceptors. Each photoreceptor represents a pixel (for the lack of a proper name). In retinal prosthesis devices, scientists use electrodes to communicate electrical signals to the optic nerves. Given the current technological limitations, it is quite challenging to fabricate a prosthesis device with 1000 electrodes (which is the minimal number of electrodes that is thought to be necessary to perform advanced visual processing [10]).
7) The prosthesis device interface that will stimulate the remaining functional part of the human retina must be convex to conform to the convex shape of the retina.
8) The signal amplitude should be variable from patient to patient. Retinal tissue resistance will vary from person to person, and this implies that the device should accommodate increasing or decreasing the electrical signal amplitude. In [4], the authors suggested that the signal frequency and shape should be variable as well as the signal amplitude, which creates a challenge for the designer.
9) The human retina processes the visual scene information to enhance the image, and possibly to extract useful and variant information from the scene. In various layers, the human retina performs spatial averaging, contrast detection, adaptation and more. Creating a retinal prosthesis device should process the visual scene in a similar way.
10) A chip that can be implanted with a smaller surgical cut is highly preferred. Engineering a chip that requires the surgeons to create a large surgical cut might be considered risky, and may not implantable.

All these considerations are important, and we are taking them into account in our proposed architecture.

4. FUTURE ARTIFICIAL RETINA ARCHITECTURE

We propose a future bio-inspired multi-layer subretinal architecture for an artificial retina. The architecture is divided into three main layers. Each layer is fabricated on its own substrate, and these substrates communicate with each other using 3D connections. Each layer is designed to have its own appropriate communication network within its cells. We propose using wireless communication, point-to-point, and other suitable communication techniques for various cells in this architecture. The three substrates are stacked above each other in a 3D fashion just like the layers in the human retina. This 3D design helps to reduce the complication of communication problems among various cells, and layers. We consider the architecture bio-inspired since it mimics the human retina cells in terms of their connectivity, and to some extent their activity. This implies that this future prosthesis device would produce an output similar to the human retina output, and it would perform similar visual scene processing.

In this section, we will describe each layer's architecture including the building blocks of the layer, and a suitable means for communications among each layer's cells and between the layers. We will roughly estimate size, and power of the architecture for the circuits in our architecture. Finally, we mention a way to implement plasticity in our architecture.

In our research into the architecture of an artificial retina, we do not attempt to solve every problem regarding retinal prosthesis devices, especially the interfaces to the optic nerve that are beyond the scope of this project. However we will give our thoughts about approaching each problem, and in some cases we will reference other researchers' solutions. The followings are a few problems that we will not tackle:

1) We are not including the electronics necessary to produce tunable output signals in terms of amplitude, and possibly frequency and signal shape. Each blind patient's retinal tissue has different resistance, which means that the retinal prosthesis device must deliver signals at different amplitudes for each client. Furthermore, some researchers [4] believe that the signal produced by ganglion cells varies in terms of frequency and shape. If that is the case, the retinal prosthesis device must be able adjust these values, which presents a quite challenging design problem. The ultimate retinal prosthesis device must be able to tune itself, be tuned prior to insertion or allow its user to tune it to produce an electrical signal that produces the best vision for each user.

2) The massive parallelism is an open design issue. Even though we are considering using nanotechnologies, we still do not know if 100 million photoreceptors can be placed on the surface of a small substrate. Each cone in the fovea is about 1.5 microns wide [16]. Carbon nanotubes that have demonstrated photoreceptor properties are about 1-2 nm in diameter [17]. If the nanotubes are arrayed perpendicularly to the surface of the substrate (standing vertically), even if we space them several diameters apart, a square array of 100 million nanotube photoreceptors would have area less than .01 mm^2 and each side would be less than .01 mm. Of course, if the nanotubes were placed so the length of the nanotube was parallel to the surface of the substrate (lying horizontally on the substrate), then the space occupied by the nanotube photoreceptor would be somewhat comparable to that occupied by a foveal cone.
3) The future retina we are envisioning would be designed to restore vision for patients who have retinal disorders. This future prosthesis device would reproduce the output of the ganglion cells in the human retina, and then send the output to the optic nerve. This architecture would not be helpful for any patient with optic nerve disorders or visual cortex problems, or any non-retinal disorder.

The ultimate purpose of a future retinal prosthesis device would be to produce an output similar to the output that is generated by the human retina. To reproduce the ganglion cells' outputs using an electronic chip, we could adopt one of two fundamentally different approaches. The first approach, the black box approach, would be to determine the output signal amplitude, shape and frequency for a set of visual scenes. Then, one would devise a mathematical function to produce this output signal given the particular visual scene information as an input. The visual scene information would include the shapes in the scene, colors of every object, illumination, and other features. This would require many experiments connecting a large number of electrodes with the output of various ganglion cells while the patient would view various scenes and then record the results. One would then need to develop a circuit that transformed the input for this scene into this output, without considering any details of the cell activity or connectivity with other cells. This methodology is not bio-inspired, and seems very tedious and probably not feasible, due to the obvious experimental hurdles, including the difficulty of placing many electrodes to capture the human ganglion cells' outputs.

The second approach would be to understand the structure of human retina well in terms of connectivity among various retinal cells, and the activity of each cell. Theoretically, if one could reproduce the connectivity of each cell with others, and produce the output that each individual cell produces, one should be able to reproduce the ganglion cells' outputs. We are approaching the retinal prosthesis architecture using this second approach. The human retina one of the most studied parts of the brain, and we know a great deal about it's neural behavior and connectivity. However, there is still much missing information about the functionality and connectivity of some cells that would have to be researched further before the ultimate retinal prosthesis would be possible. In order to be fully functional, the architecture we are envisioning should capture what is known about neural behavior and connectivity.

When proposing a design for retinal prosthesis, the first question raised is whether the device would be completely inside the eye (subretinal) or partially inside the eye (epiretinal). Most researchers aiming to produce a prosthesis device commercially in the next very few years minimize the implantable part of the device. The current technologies limit the number of implantable electrodes and there is a limit in point-to-point wiring capacity. As such, minimizing the implantable part is a desirable approach for retinal prosthesis designers who are building devices at the present time, and thus an epiretinal architecture may be more suitable. However, we are envisioning an architecture that will be feasible in the future decades when we have more advanced nanotechnological capability, and better on-chip communication technologies. Furthermore, we believe that subretinal devices would be more convenient for patients since there would not be any external parts of the devices, also removing the communication problem between the internal and external modules.

Since we are targeting technologies available in the future, and because it is more convenient for patients not to have external parts for a prosthetic device, we have decided to investigate the feasibility of a subretinal architecture (fully implanted inside the eye). A subretinal architecture would be more challenging than an epiretinal architecture in terms of power source, generated heat, and size since the whole device would be implanted in the eye.

4.1 Powering the Future Retinal Implant

Device power is always a consideration with implanted electronics. There are four major options for power sources. The first option is to use a battery. Having a battery inside the eye is quite inconvenient for the patient, as surgeons would

have to perform eye surgery every time the battery must be replaced. We believe using a battery would only be desirable if it were possible to recharge the battery inside the eye without surgery. The second option is to use body heat. A thermoelectric generator is needed to convert the body heat to potential that could be used to power the prosthetic device. Even though we did not investigate the technological issues regarding using body temperature, it might be a feasible option as that approach is being using in other prosthetic devices. The third option is to transmit wireless power to the subretinal device from an external source using power telemetry. Using an external system may not be convenient for the patient, but this is an option that has been shown feasible in the DOE artificial retina. The fourth option is the use of incident light. When the incident light falls into the device, it provides a power source for a subretinal chip. The Optobionics team used incident light to power their ASR chip. In this approach, the chip needs to be able to operate from the generated potential regardless of the incident light intensity. That is, given a scene with low illumination, the chip should still be able to generate enough electrical power to operate at reduced capability, just as the biological eye does.

Relying on incident light would require high-quality light transduction circuitry in our first layer that is capable of absorbing the incident light. Furthermore, this circuitry should be able to detect the entire or majority of the visible spectrum. In the discussion of layer one of our proposed architecture, we will tradeoff various options for light transduction including regular photodiodes, PNP transistors and carbon nanotubes.

The heat generated as power is consumed when operating a subretinal device poses a critical engineering challenge for any retinal prosthesis implant. The human retina uses a tiny amount of power and hence generates a small amount of heat. If the implanted device were to produce more heat than the human retina, it could injure remaining healthy eye tissue including the optic nerve, and possibly lead to irreversible vision damage.

We will estimate the power consumed for certain components in our architecture, and draw some conclusions about future feasibilities as a result of this estimate. The major power-dissipating circuit in the neurons we envision is expected to be the analog adders that sum potentials because there is a flow of current from the power supply to ground regardless of the inputs to the adders [15]. For every input to each neuron (synapse or gap junction) there would be approximately one two-input adder. The adders we have designed and simulated add excitatory post-synaptic potentials (EPSPs) as positive potentials and inhibitory post-synaptic potentials (IPSPs) as negative potentials, essentially subtracting them. The adder circuits use carbon nanotubes as transistor circuit elements and the circuits have been simulated with SPICE models [22]. The simulations showed that the average power consumption ranged from 85.1 μW to 87.5 μW averaged over an 80 ps window, depending on the potential amplitudes being added. The synapses and axon hillock for spiking neurons are not expected to contribute appreciably to this amount since there is no DC current to ground. To achieve biological levels of potentials, the voltages in the adder circuits would have to be scaled down by at least a factor of 10, reducing the maximum power consumption per adder to <8.75 μW, since power would scale down more than a factor of 10, possibly by a factor of as much as 100. The total power consumption would then depend on the number of inputs to each neuron to be summed.[2] Assuming 100 million photodetectors producing voltages to be summed in some fashion, at least 100 million adders would be required in Layer I of an artificial retina, resulting in a power consumption of between 87.5 and 875.0 W. Since these power requirements are clearly excessive, the number of photoreceptors in the fully-functioning retina would have to be reduced. Given the power requirements of each neuron in the retina, we expect that a fully-functioning retina with any sort of summation of potentials would be limited to between 12,000 and 120,000 photoreceptors using our proposed adders.

Thus it appears unlikely using our current neural designs that a subretinal prosthesis could achieve the processing power of the human retina unless economies of scale and coverage were used outside the fovea, and/or the adder circuit was replaced with a more basic low-power circuit that summed potentials to replace the power-hungry adders. A lower-power adder would be invaluable for this application.

4.2 Layers in the Architecture of the Future Artificial Retina

Our architecture consists of three layers of processing, as shown in Figure 1. Layer I contains photoreceptors and horizontal cells. Layer II consists of bipolar and amacrine cells. Finally, Layer III consists of ganglion cells. When light falls into the eye, it will fall onto layer I. Layer one will convert the light into electrical signals, and will provide some

[2] It should be noted that the additions and subtractions of potentials occur in a distributed fashion and it is possible that the structure of the dendritic arbor would affect how such arithmetic is carried out in each neuron, as in the cortex [18].

functionality such as spatial averaging and contrast enhancement through the horizontal cells. Then, the signal will go to layer II which in turn processes the photoreceptors' outputs, and sends the signal to Layer III. Layer III contains 1/10 the number of cells as layer I's photoreceptor cells, just like the human retina. Layer III's main's task is to stimulate the remaining part of the human retina. Layer I has the largest number of cells, and hence occupies the largest volume. Layer III has the smallest volume, and it has the curvature that will help to position the artificial prosthetic device.

In layer one, we have photoreceptors and horizontal cells. This layer would convert light into signals, and would be responsible for contrast sensitivity and normalization due to overall illumination differences.

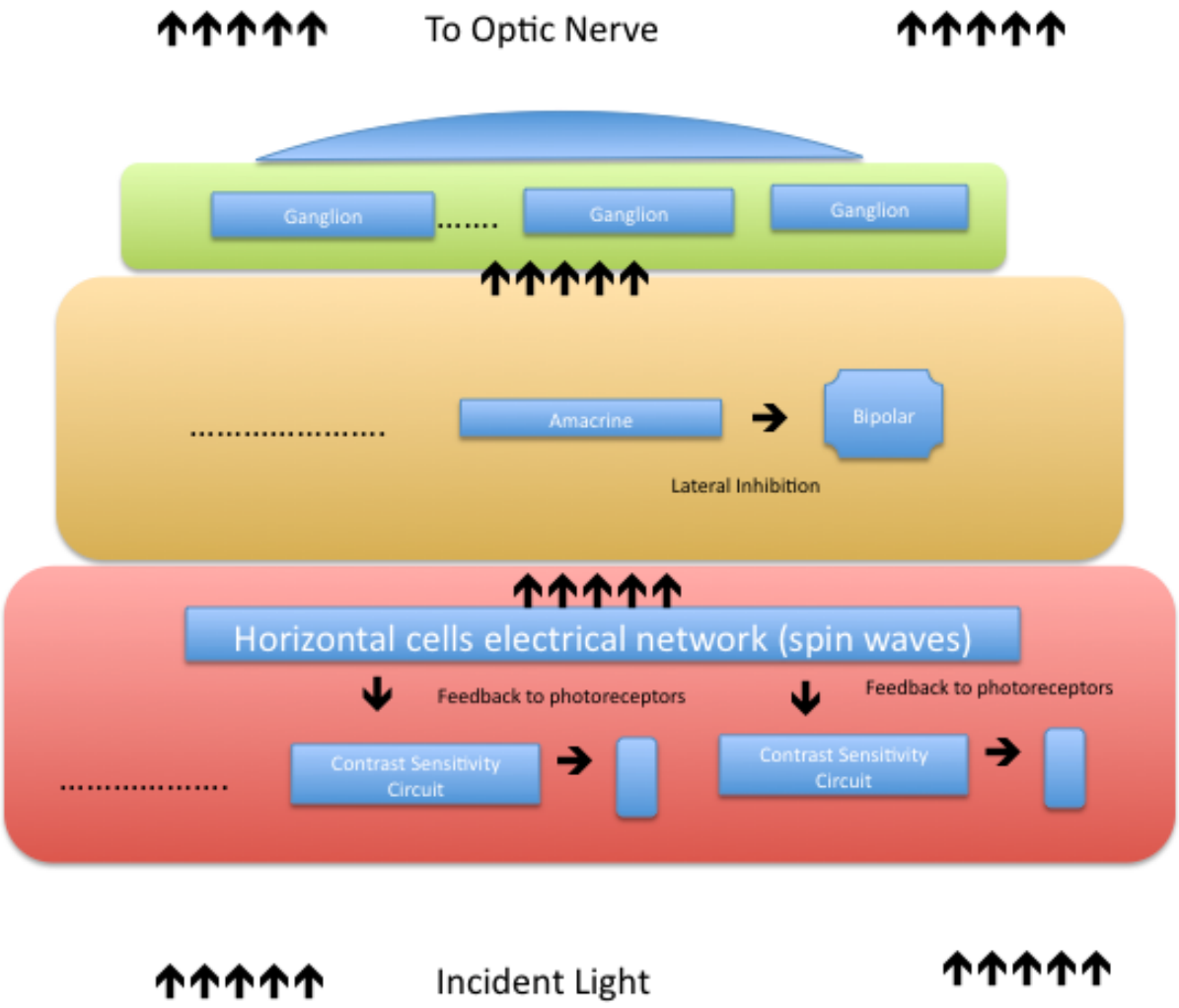

Figure 1: The Proposed Retinal Architecture

For light transduction, we could use photodiodes, Bipolar Junction Transistors (BJT) or carbon nanotubes. Photodiodes and BJTs operate in a quite similar way. Once a photon bombards the silicon of a photodiode or a BJT, it creates a hole, and current flows. Both Photodiodes and BJTs have been used for light transduction for long time in many fields including artificial retinas [6, 7]. There are no real shortcomings with using these technologies except for size and power consumption compared to carbon nanotubes. Recently, researchers from Sandia National Laboratories have created the first carbon-nanotube devices that can detect the entire visible spectrum of light [5], as shown in Figure 2. These nanotubes are decorated with three kinds of chromophore molecules (red, green, and blue) that respond to each

wavelength by changing their shapes. This will eventually change the electrical conductivity of the nanotube in a way that can be measured to deduce the color and intensity of the light. Their work is still in its early stages but it sounds very promising for the artificial prosthesis field.

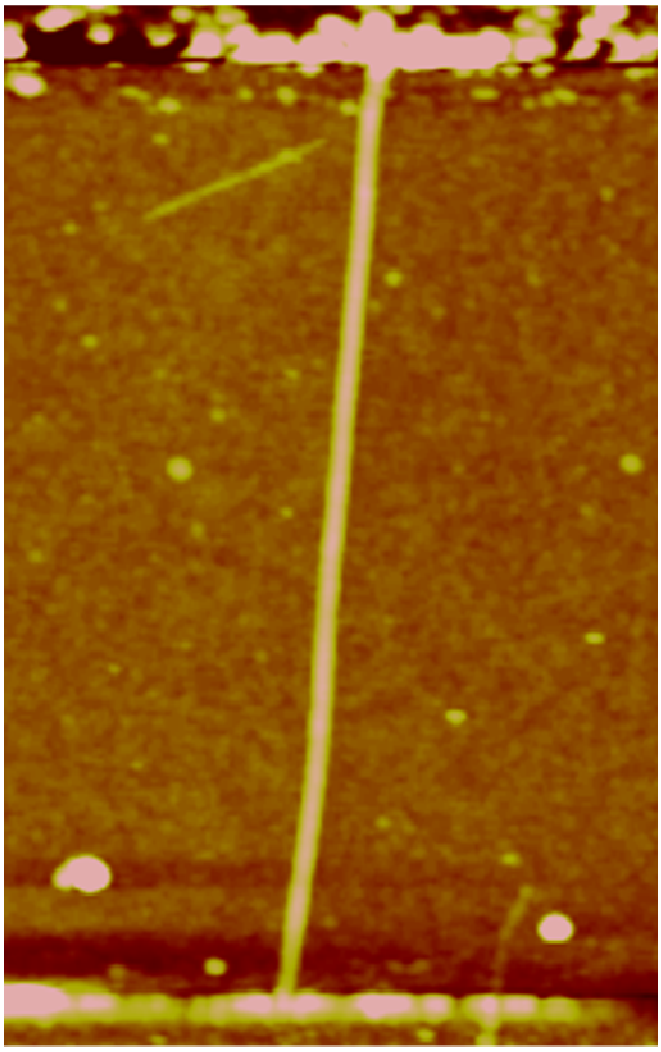

Figure 2: A photosensitive nanotube [17] (Credit: Xinjian Zhou, Sandia)

We propose the usage of spin waves (magnetic nanowaves) to emulate the communication among the horizontal cells [24,26]. An electron's spin produces a magnetic moment, and if a magnetic field is applied, this spin moment will precess about the field. A collection of such spin precessions is a spin wave. These magnetic waves propagate along the surface of a substrate. Spin waves are suitable since the technology helps reduce the complexity of both the power and also the interconnection problem at the same time. The spin wave medium does not require any physical connections (no need for point to point or network-on-chip connections, for example), which reduces the complexity of the interconnectivity problem. Regarding power consumption, spin waves do not transport on the medium using particles (holes or electrons). Particles would encounter dissipative resistance once they transported in the medium, and the higher the resistance, the higher the power consumed. Since spin waves do not transport using particles, they will have less power consumed and less heat generated, which is very suitable for an artificial retina. Furthermore, spin waves can travel only on the surface rather than all directions (e.g. radio waves). In earlier architectures, Mahowald *et al*. proposed using a resistive network to emulate the horizontal network electrical network [7]. In their design, each node in the

horizontal cell layer connected to 6 neighboring nodes using resistors. This resistive network performed spatial averaging that is done by the horizontal cell layer. Boahen *et al*. proposed using MOS transistors working in the subthreshold region to connect the horizontal cells [6].

In the design we propose, each photoreceptor is connected to a spin wave transmitter that transmits magnetic waves on the surface that represents the horizontal layer electrical network. This surface represents each horizontal cell's electrical network and communications between horizontal cells is performed by these generated spin waves. At any point in this surface, the spin waves have a certain strength that depends on whether the nearby photoreceptors are depolarizing or not. Furthermore, every transmitter circuit is associated with a spin wave receiver that measures the strength of spin waves at that point. We would feed the outputs of the photodiode and spin wave receiver to a post-processing circuit. The output of this circuit would be connected to the second layer.

Layer II receives its input from the post-processing circuit output. The basic operation of layer II is to process layer I output and pass it to Layer III. Assuming we have N Ganglion cells, we will have about 100N photoreceptors. Layer II is responsible for combining about 10 photoreceptor outputs into one output. In Layer II, we have two major types of cells: bipolar cells, and amacrine cells. Amacrine cells perform lateral inhibition on bipolar cells, along with other functions including edge detection. It is useful to group the amacrine cell types into the general descriptors of narrow-field (30-150um), small-field (150-300 um), medium-field (300-500 um) and wide-field (>500 um) based on a measurement of their dendritic field diameters [25]. We propose using spin wave transmission for some amacrine cells to achieve lateral inhibition. However, the amacrine cells' dendritic field diameters vary, some amacrine cells overlap and some tile the surface of the inner plexiform layer. Given the variations in amacrine cell communication, we need to have spin wave transmitters with differing transmission strength, and with different communication architectures.

Processing in layer II could be carried out by circuits constructed of carbon nanotube transistors. Our collaborator, Chongwu Zhou at USC, has constructed logic gates with nanotubes fabricated by his group, as shown in Figure 3. Laboratory construction of simplified neural circuits is underway.

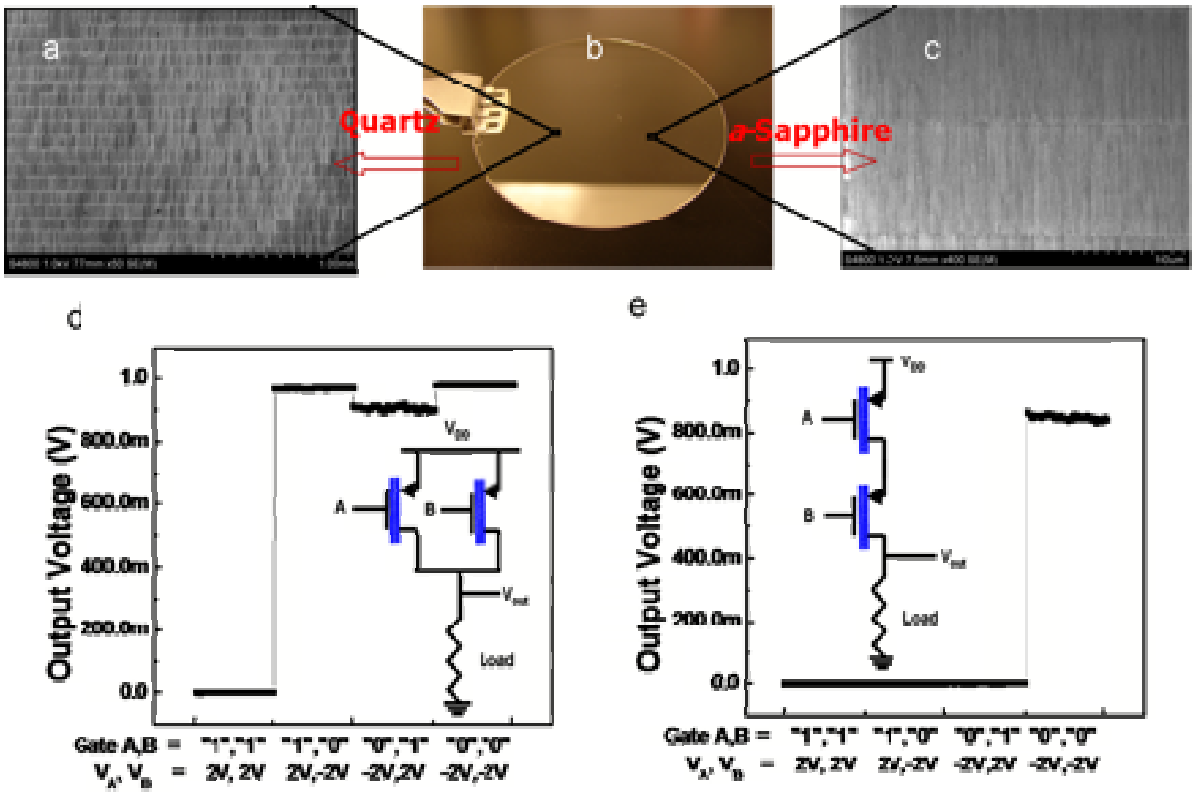

Figure 3: a), c) SEM image of aligned nanotube arrays on quartz and sapphire, respectively. Figure b) photograph of 4 inch substrate with the as-grown nanotubes. Figure d) The NAND logic device data. Inset: schematic diagram of NAND device. Figure e) The NOR logice device data. Inset: schematic diagram of NOR device

Layer III consists of ganglion cells. Basically, this layer receives layer II output and stimulates the optic nerve axons. This layer should be convex in shape to allow surgeons to position the implant in the back of the eye. The optic nerve will communicate spikes to the human brain.

Our estimates for the volume occupied by the retinal cells is very loose since the technology to assemble carbon nanotube circuits as complex as the neurons we envision is yet to be determined. However, we have simulated a variety of neural structures using SPICE simulation models of carbon nanotube transistors. A basic synapse circuit is shown in Figure 4 a). Entire neurons (including dendritic computations, and excitatory and inhibitory synapses) have been simulated. Figure 4 b) (credit: Khushood Irani) shows our conception of a 3-dimensional carbon nanotube synapse. We base our retinal volume estimates on this particular structure. We assume that the nanotubes are about 100 nm in length and 2 nm in diameter. The horizontal nanotubes are semiconducting, and the vertical ones act as metallic connections. The flat material is graphene and is a single molecule thick.

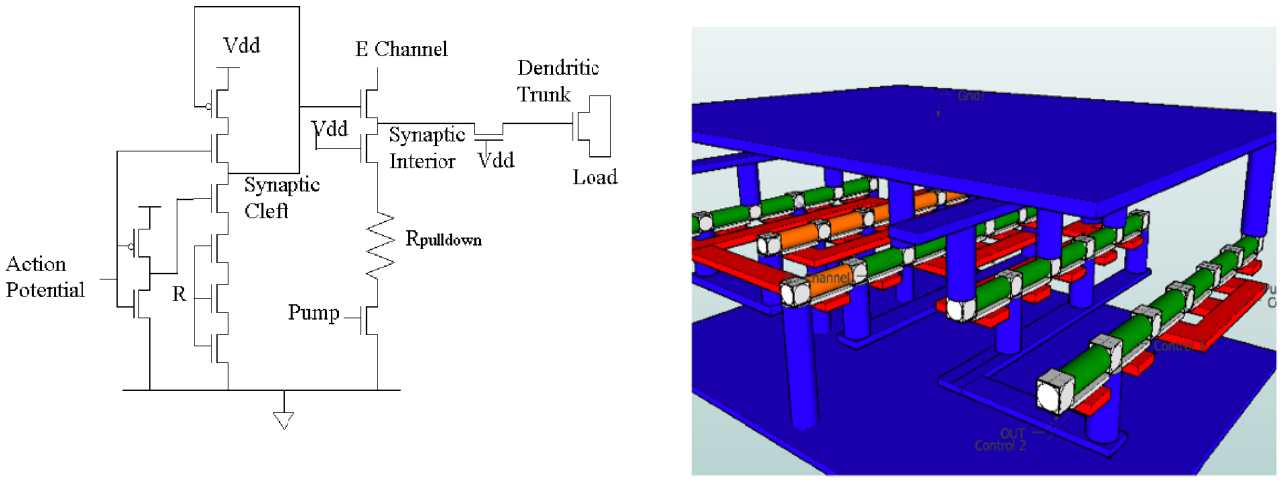

Figure 4: a) The basic synapse circuit used for our volume estimations. b) A 3D model of the basic synapse

We have begun designing synapse circuits that exhibit plasticity. The current plasticity mechanism we envision is going to increase the synaptic complexity significantly, maybe as much as a factor of 4. These results will be published elsewhere. We envision that processing delays will not be problematic with the circuits we envision, since nanotube circuits exhibit picosecond delays, and the signals would actually need to be delayed substantially to match the biological signal timing.

5. CONCLUSION

We propose a future bio-inspired multi-layer subretinal architecture for an artificial retina. We mention a few retinal architecture challenges. We discuss few power supply, and photo transduction options for our retinal architecture. We consider using carbon nanotubes as an alternative technology for the retinal photoreceptors. We calculate loose estimates for power, size and delay. We show it is possible to have millions of carbon nanotube photoreceptors in a small area, but the predicted power consumption for a full-size retina is problematic.

REFERENCES

[1] Margalit, E., Maia, M., Weiland, J., Greenberg, R., Fujii, G., Torres, G., Piyathaisere, D., Ohearn, T., Liu, W., and Lazzi, G., "Retinal prosthesis for the blind," Survey of Ophthalmology, 47(4), 335-356 (2002).

[2] Brindley, G. S. and Lewin, W. S., "The sensations produced by electrical stimulation of the visual cortex," J Physiol, 196(2), 479-493 (1968).

[3] Chow, A. Y., Chow, V. Y., Packo, K. H., Pollack, J. S., Peyman, G. A., and Schuchard, R., "The artificial silicon retina microchip for the treatment of vision loss from retinitis pigmentosa," Archives of ophthalmology, 122 (4), 460-469 (2004).

[4] Scribner, D., Johnson, L., Klein, R., Bassett, W., Howard, J. G., Skeath, P., Wasserman, L., Wright, B., Perkins, F., Peckerar, M., Finch, B. J., Graham, R., Trautfield, C., Taylor, S., and Humayun, M., "A Retinal Prosthesis Device: Based on an 80x40 Hybrid Microelectronic-Microwire Glass Array," Custom Integrated Circuits Conference, Proc. IEEE, 517-520 (2003).

[5] http://www.technologyreview.com/computing/22257/?nlid=1837&a=f

[6] Boahen, K. A., and Andreou, A. G., "A Contrast sensitive Silicon Retina with Reciprocal Synapses," Advances in neural information processing systems, 4, 764-772 (1992).

[7] Mahowald, M., "VLSI Analogs of Neuronal Visual Processing: A Synthesis of Form and Function, " Phd dissertation at the California Institute Of Technology (1992).

[8] Zaghloul K. A, "A silicon implementation of a novel model for retinal processing, " Phd dissertation at the University of Pennsylvania (2001).

[9] Maynard, E. M., Nordhausen, C. T., and Normann, R. A., "The Utah Intracortical Electrode Array: A recording structure for potential brain-computer interfaces, " Electroencephalography and clinical neurophysiology 102 (3), 228–239 (1997).

[10] Weiland, J. D., Liu, W., and Humayun, M. S., "Retinal Prosthesis annual review, " Annual Review of Biomedical Engineering, 7(1), 361–401 (2005).

[11] Wyatt, J. L., Standley, D. L., and Yang, W., "The MIT Vision Chip Project: Analog VLSI Systems for Fast Image Acquisition and Early Vision Processing, " Robotics and Automation, 2, 1330–1335 (1991).

[12] Rizzo, J., Loewenstein, J., and Wyatt, J., [Retinal Degenerative Diseases and Experimental Therapy], Springer US, 463-469 (1999).

[13] Rizzo, J. F., Wyatt, J., Loewenstein, J., Kelly, S., and Shire, D., "Perceptual efficacy of electrical stimulation of human retina with a microelectrode array during short-term surgical trials, " Invest. Ophthalmol. Vis. Sci. 44 (12), 5362–5369 (2003).

[14] Eckmiller, R., "Learning retina implants with epiretinal contacts, " Ophthalmic research 29 (5), 281–289. (1997).

[15] Parker, A. C., Joshi, J., Hsu, C.-C., Singh, N. A. D., "A carbon nanotube implementation of temporal and spatial dendritic computations, " Circuits and Systems, MWSCAS 2008, 818–821 (2008).

[16] http://webvision.med.utah.edu/photo1.html

[17] Zhou, X., Zifer, T., Wong, B. M., Krafcik, K. L., Leonard, F., Vance, A. L., "Color detection using chromophore-nanotube hybrid devices, " Nano Letters, 9 (3), 1028–1033 (2009).

[18] Polsky, A., Mel, B., Schiller, J., “Computational subunits in thin dendrites of pyramidal cells,” Nature Neuroscience, 7, 621 - 627 (2004).

[19] https://www.stanford.edu/group/brainsinsilicon/vision.html

[20] Moravec, H., “When will computer hardware match the human brain?” Journal of Evolution and Technology, 1 (1998).

[21] Moravec, H., [Robot: Mere Machine to Transcendent Mind], Oxford University Press, USA, Chapter 3 (2000).

[22] Deng, J., Wong, H. S. P., "A circuit-compatible spice model for enhancement mode carbon nanotube field effect transistors, " Simulation of Semiconductor Processes and Devices, pp. 166–169 (2006).

[23] Silva, G., A., "Neuroscience nanotechnology: progress, opportunities and challenges, " Nature Reviews Neuroscience, 7, 65-74 (2006).

[24] Eshaghian, M. M., Khitun, A., Navab, S., Wang, K., “A Nanoscale Crossbar with Spin Waves,” Proc. of the 6th IEEE Conference on Nanotechnology, 1, 326- 329 (2006).

[25] Kolb, H., Nelson, R. and Mariani, A., “Amacrine cells, bipolar cells and ganglion cells of the cat retina: A Golgi study, “ Vision Res., 21, 1081-1114 (1981).

[26] Eshaghian, M. M., Friesz, A. K., Khitun, A., Navab, S., Parker, A. C., Wang, K. L., and Zhou, C., “Emulation of Neural Networks on a Nanoscale Architecture,” special issue of the Journal of Physics: Conference Series, 61, 288-292 (2007).