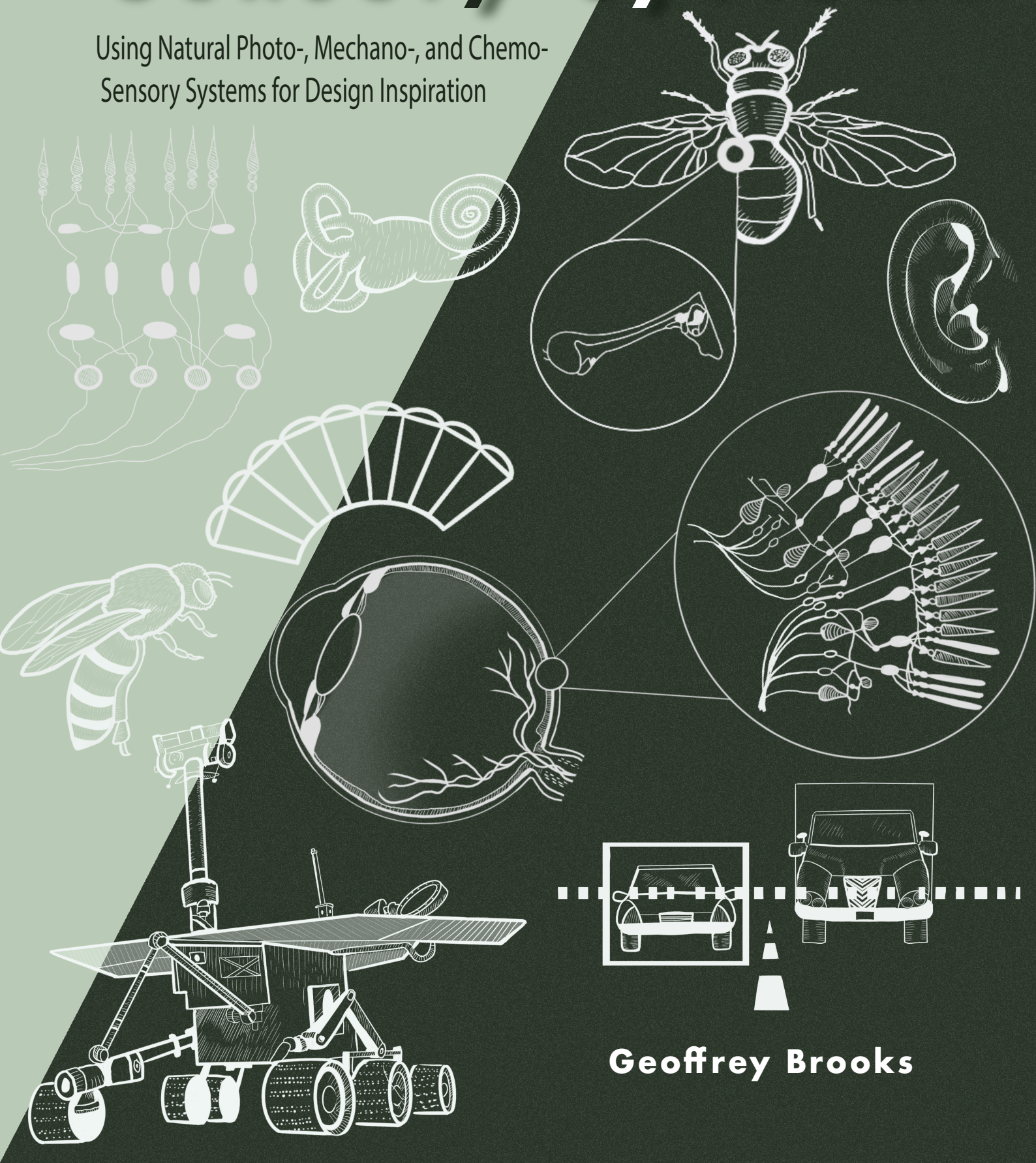
Bio- Inspired Sensory Systems

Using Natural Photo-, Mechano-, and Chemo-Sensory Systems for Design Inspiration



Geoffrey Brooks

Bio-Inspired Sensory Systems

Using Natural Photo-, Mechano-, and Chemo-Sensory Systems for Design Inspiration

Dr. Geoffrey Brooks

This work is published by Florida State Open Publishing, 116 Honors Way, Tallahassee, FL 32306.

ISBN: 978-1-7365779-0-5
DOI: https://doi.org/10.33009/fsop_brooks0121

In the name of my Savior, the Lord Jesus Christ,
With thanks to God the Father
Soli Deo Gloria

"Whatever you do, whether in word
or deed, do it all in the name of
the Lord Jesus, giving thanks to
God the Father through Him"
Colossians 3:17

Contents

Preface

Mimicking biology to improve sensory system designs and signal processing algorithms has thrived in the past and will continue to do so for decades to come. Technological advances have generally followed Moore's Law (capability doubling every 2 years or so) since the 1960's while our understanding of biological sensory systems is also rapidly advancing. These trends fuel the fertile grounds of bio-inspired sensory systems, a topic that is inherently multidisciplinary. This book is intended to be used as the primary text for a technical elective course in an undergraduate or graduate electrical engineering curriculum, but it could certainly be used for related purposes as well. Available student materials for such a course have previously been limited to biology sensory system texts, robotic application systems, collections of papers from numerous authors, current technical publications, and related material that have been useful but awkward as student study material due to the complexity of biology and the vast array of technical applications. There has not been a book or summary of study materials available that systematically covers natural photo-, mechano-, and chemo-sensory systems across the animal kingdom and also summarizes various novel engineering ideas that glean ideas from these natural sensory systems.

For a one-semester course it is expected that the instructor of such a course would search the literature for recent sensory system designs inspired by biology such as examples already covered in this text . It is recommended that special assignments be given to students in the course to review the literature for relevant papers (or be given a set of papers to choose from) and present their findings to the class. A suggested assignment is given in Appendix A. It is also recommended that the instructor include the following free texts for more in-depth coverage of specific linear systems theory and image processing concepts when applicable (for example, when discussing a particular current application):

Ulaby, F. and Yagle, A, *Signals and Systems: Theory and Applications*, Michigan Publishing, ISBN 978-1-60785-487-6, 2018. Available at ss2.eecs.umich.edu.

Yagle, A. and Ulaby, F., *Image Processing for Engineers*, Michigan Publishing, ISBN 978-1-60785-489-0, 2018. Available at ip.eecs.umich.edu.

The following text is referenced frequently and is recommended for a much more thorough study of the structure and function of natural sensory systems. Breaking this topic into photo-, mechano- and chemo-sensory systems is inspired by the organization of this book:

Smith, C. U. M., Biology of Sensory Systems, 2/e, John Wiley, and Sons, ISBN: 978-0-470-51862-5, 2008.

Since solutions manuals to textbooks are so readily available, it makes more sense to work *example* problems in this text and provide similar *exercise* problems (with only the answers provided) to assess the skill required to solve the problems. There are also sets of *questions* for the student to check their comprehension of the material covered. Most of these questions are directly answered in the text so an answer key is not provided separately.

Specific problems are introduced and worked that are intended to reinforce specific concepts covered. There are many various other problems that could be introduced, but the goal is to keep within the scope of a one-semester course. The following is a partial list of the problem types and why they were chosen for this text:

2D convolution problems show how image filtering is a 2D extension of 1D convolution covered in a standard linear systems course.

Space constant problem shows significant attenuation of ionic signal as it travels down the axon.

Neuronal circuit model problems emphasize that many times neuronal electronic signals are due to the movement of ions and not electrons and holes. In biology ions must be replenished; thus, the circuit includes dependent sources that model microbiological structures called *ion pumps.*

Motion detection problems show that delayed responses of adjacent neurons are needed for most basic motion detection queues.

Center-surround opponent processing problems show how three cone types with broadly overlapping responses across the visible electromagnetic spectrum can be combined to uniquely identify very specific colors.

Wavelet analysis and synthesis problem demonstrate that broadly overlapping filters can be used to encode

detailed signals and conserve signal energy, which is very important to biology.

Auditory neuron response problem is another example of how broadly overlapping responses of adjacent neurons can be used to extract specific tonal information from incoming sound source.

The author is grateful for a one-semester professional development leave (PDL) assignment in 2020 that made possible the completion of this project. Randy Hanna, Dean of Florida State University Panama City (FSU PC), is appreciated for his willingness to inaugurate PDL opportunities for dedicated teaching faculty on our campus. Shaun Saxon, FSU PC Librarian, and Laura Miller, FSU Open Publishing Librarian were invaluable for the help and advice for making this available in the most practical sense. Betul Adalier is very much appreciated for sharing her design talent in creating the front and back cover.

This book is provided for free in accordance with the Creative Commons license stated earlier. It is requested that you let us know how you plan to use the book and to let us know how we can make it better. The author may be contacted directly for comments and feedback at gbrooks@fsu.edu. The book may be downloaded from the FSU Libraries at https://manifold.lib.fsu.edu/projects/bio-inspired-sensory-systems.

Chapter 1: INTRODUCTION

As an introduction we will define some multidisciplinary terminology, consider our motivations, and cover some relevant academic activity as well as research publications.

1.1 Relevant terminology and related bio-inspired technologies

Biomimetic implies the mimicry of biology; note the word *mime* embedded in the term. In recent years *biomimicry* seems to be used more and more for sensory systems applications (our interest), while *biomimetics* implies molecular-level mimicry. This text is focused on biologically-inspired paradigms used for sensory systems and the signal processing that goes with such systems. The subject is sensory systems and not the research associated with mimicking organic chemistry, muscle tissue, etc. In this text cursory descriptions of biological phenomena will be followed by electronic sensor designs and the signal processing algorithms that emulating such phenomena for useful technological application. *Bioprincipic* is a similar term used recently implying the mimicry of biological principles.

Biometric implies measuring biological features unique to an individual to determine the identity of the person. For example, authentication can be granted based on a pattern matching of a fingerprint, scanned iris image, or recorded voice pattern. This could be used for building and computer security purposes.

Biomedical means the branch of medicine associated with survival in stressing environments. *Bionic* means enhancing normal biological capability with electronic or mechanical devices [Webster].

Bioinformatics is used to describe computer applications of extracting information *about* biological phenomena, primarily in the field of molecular biology. Bioinformatics is more formally defined as " The collection, classification, storage, and analysis of biochemical and biological information using computers especially as applied to molecular genetics and genomics." [Webster].

Anatomy and *Physiology* imply structure and function, respectively. Scientists from many disciplines often organize their thoughts in similar ways. However, until there is a reason to communicate across disciplines, the terminology in each tends to develop into quite different terms. Table 1 is an observation of the separation of phenomena into physical and abstract categories.

Genetic Algorithms refers to computational methods inspired by genetics. A genetic algorithm may consist of randomly choosing a solution to a problem or improving an existing solution based on an evaluation of fitness representing the problem solution. Improved solutions may be derived from fitness evaluation and genetic operators representing mutation and crossover.

Evolutionary Computation refers to the computational methods mimicking natural evolutionary forces.

Neural Networks is used to refer to networks of computational elements that process information in an analogous way to biological neuronal networks.

Both natural and artificial neural networks perform a nonlinear transform on an aggregation of many weighted input signals. There are many artificial "neural network" paradigms (ANN's) that include many ideas ont found in biological neuronal networks, although the general concept has its original inspiration from biology.

Table 1. Different Terminology, Similar Concepts

Concept	Biology	Engineering	Computer	Common Term
Physical	Anatomy	Architecture	Hardware	Structure
Abstract	Physiology	Algorithm	Software	Function

Nevertheless, most ANN variations have these features in common with natural neural networks:

- A summation of many inputs, each weighted differently based on learned examples
- A non-linear output mapping function follows the summation
- massively parallel
- distributive processing
- adaptive

The application of ANN's to various engineering applications has grown into an academic field of its own, with separate texts and courses dedicated to the field. Further study of neural networks is reserved for courses and texts dedicated to this subject.

1.2 Motivation for this multidisciplinary study

So, why a special electrical engineering course focused on biologically-inspired sensory system designs and signal processing techniques? A few of the reasons include:

- Natural systems solve engineering problems
- Biological information is becoming increasingly more available
- Technology is becoming increasingly more affordable and available
- Research agencies continue to support bio-inspiration
- A better understanding of biology can result from attempting to imitate biology

1.2.1 Natural systems solve engineering problems

From the earliest times we have looked to biological systems for engineering solutions to our technical problems. For example, in Greek mythology the legendary Daedalus, builder of the Cretan labyrinth, was motivated by birds to build wings to help him and his son, Icarus, escape imprisonment. Later observations of birds, such as wing-shape, have led to modern aircraft design features.

Velcro has been inspired by the way burrs attach themselves to clothing. Autonomous robots can benefit from the study of natural control mechanisms found in similar creatures in the animal kingdom. Machine vision systems for robotics require the separation of objects from the background, a task inherently embedded into the design of natural vision systems. The image recognition capability of humans is difficult to duplicate with computer technology, although neurons are five or six orders of magnitude slower than silicon transistors and heterogeneous (or considerably 'mismatched' when compared to transistors).

1.2.2 Biological information is becoming increasingly more available

The difficulty in reverse-engineering natural systems is due in part to our lack of complete understanding of these complex systems. In organic chemistry and microbiology, we have uncovered much detail of the fundamental physical processes at the neuronal level. We also have considerable understanding of the overall systems behavior from fields such as psychology or psycho-physics. What is difficult to grasp, however, is how the microscopic processes transforms sensory information into the macroscopic decisions and behaviors. This leads to an interest in natural design optimizations and interconnection schemes.

It is commonly agreed that many people will do almost anything for money but will also freely give it up for their health. This captures our limited existence in time and space while desiring permanence, which leads to our willingness to do whatever we can to maintain or improve our health. As a result, there is and will always be enormous resources (funds, etc.) available for exploring a deeper understanding of biological phenomena. Although guided for medical purposes, system concepts applicable for other uses will eventually unfold. As we move further into the information age with better and better technology, many of the details are already available for exploiting natural sensory design concepts.

Although there is already an abundance of information available on natural sensory systems and signal processing, it is difficult for engineers to decipher useful information from the biomedical literature. This is due in part to the different motivations: The medical community is interested in diagnosing (organic) system problems and formulating procedures and medications to fix those problems or allow the patient the ability to adequately deal with the problems. The engineer, on the other hand, is more interested in how specific tasks are accomplished from the available sensory signals.

1.2.3 Abundant technology is affordable and user-friendly

Due to rapid advances in processing speeds and throughput capabilities many successful applications have now been developed using artificial intelligence, deep learning neural network architectures, and other related technologies. A small sample of tools readily available for students and researchers include:

Reconfigurable computing tools such as *Quartus* (Altera) and *Vitis* (Xilinx)

Circuit simulation tools such as *PSPICE* (Microsim)

Data Acquisition such as *LabVIEW* (National Instruments)

Computational tools such as *Matlab* (Mathworks)

Development platforms such as Raspberry, Arduino, etc.

Languages such as *Python*

1.2.4 Research agencies continue to support bio-inspiration

The author draws from former work experience at the Munitions Directorate of the Air Force Research Laboratory (AFRL/MN). To address the high signal processing throughput and short latency of an imager that guides an exo-atmospheric hypervelocity missile, novel concepts were explored that involved biologically-inspired approaches. Funded concepts included an infrared sensor with retina-inspired readout, multi-resolution targeting inspired by foveated vision, and other research projects exploiting various bio-inspired sensory design ideas.

Some historical efforts (late 1980's and 1990's)

Much of the work at AFRL/MN was leveraged from former research sponsored by the Defense Advanced Research Projects Agency (DARPA) and the Office of Naval Research (ONR). Research funded by ONR and DARPA as well as National Science Foundation (NSF), National Institute of Health (NIH), and others have resulted in books whose individual chapters are written by the various researchers, which can lead to a considerable lack continuity and consistency. Nevertheless, the material in such books is proven to very useful; a few examples include

Mead, Carver, *Analog VLSI and Neural Systems*, Addison-Wesley, 1989.

Zornetzer, Steven, Davis, Joel, and Lau, Clifford, editors, *An Introduction to Neural and Electronic Networks*, Academic Press, 1990.

Ayers, J., Davis, J. and Rudolph, A., editors *Neurotechnology for Biomimetic Robots*, MIT Press, 2002.

Bar-Cohen, Yoseph, and Breazeal, Cynthia, editors, *Biologically-inspired Intelligent Robots*, Taylor and Francis, 2003.

Bar-Cohen, Yoseph, editor, *Biomimetics: Biologically-inspired Technologies*, Taylor and Francis, 2006.

The following book and the 2nd edition have been useful for covering the structure and function biological sensory systems:

Smith, C.U.M, *Biology of Sensory Systems*, John Wiley and Sons, ISBN: 0-471-85461-1, 2000.

As an example of continued strong and direct support for biomimetics, consider this excerpt from an announcement for *Biomimetics for Computer Network Security Workshop* (1999):

> *"The Office of Naval Research is sponsoring a workshop whose goal will be to identify technologies that are inspired by biological foundation and that, when matured, may contribute to a significant increase in network security capability...This research is aimed at developing a new class of biologically inspired robots that exhibit much greater robustness in performance in unstructured environments than today's robots.... The research involves a close collaboration among robotics and physiology researchers at Stanford, U.C. Berkeley, Harvard and Johns Hopkins Universities... sponsored by the Office of Naval Research under grant N00014–98–1–0669..."*

More recent developments

In August, 2020, the Office of Naval Research (ONR, www.onr.navy.mil, Code 341) continued to solicit contract and grant proposals in the area of "Bio-inspired Autonomous Systems" with the following description:

> *The aim of Bio-inspired Autonomous Systems is to extract principles of sensorimotor control, biomechanics and fluid dynamics of underwater propulsion and control in aquatic and amphibious animals that underlie the agility, stealth, efficiency, and sensory adaptations of these animals. The principles that emerge from this interdisciplinary research are formalized and explored in advanced prototypes. The goal of this program is to expand the operational envelope of Navy underwater and amphibious vehicles and enable enhanced underwater manipulation.*

as well as in "Bio-inspired Signature Management" with the following description:

> *The Bio-inspired Signature Management program aims to discover biologically-inspired adaptations and bioengineered solutions to expand current warfighter capabilities in detection mitigation and undersea navigational challenges. This will be accomplished through multidisciplinary research in science and technology fields such as bio-inspired / biomimetic materials, visual and sensory perception, and bio-optics / bioelectronics.*

Also in August 2020 the Defense Advanced Research Projects Agency (DARPA, www.darpa.mil) gives the following description of their "Nature As Computer (NAC)" program:

> *Certain natural processes perform par excellence computation with levels of efficiency unmatched by classical digital models. Levinthal's Paradox illustrates this well: In nature, proteins fold spontaneously at short timescales (milliseconds) whereas no efficient solution exists for solving protein-folding problems using digital computing. The Nature as Computer (NAC) program proposes that in nature there is synergy between dynamics and physical constraints to accomplish effective computation with minimal resources. NAC aims to develop innovative research concepts that exploit the interplay between dynamic behaviors and intrinsic material properties to develop powerful new forms of computation. The ability to harness physical processes for purposeful computation has already been demonstrated at lab-scales. NAC seeks to apply these concepts to computation challenges that, for fundamental reasons, are poorly suited to, or functionally unexplored with, classical models. NAC will lay the foundation for advancing new theories, design concepts and tools for novel computing substrates, and develop metrics for comparing performance and utility. If successful, NAC will demonstrate the feasibility of solving challenging computation problems with orders-of-magnitude improvements over the state of the art.*

1.2.5 Imitating biology can lead to a better understanding of biology

Although engineering applications may result from biological inspiration, sometimes those applications are biomedical. For example, artificial neural networks are used for identifying potential cancerous sites in x-ray images. Meanwhile, biomimetic robots are not only used as testbeds for potential engineering applications, but also as tools for biologists to better understand complex animal-environment relationships. An example of this expressed is found concerning MIT's "RoboLobster" in the following quote:

> *The major result of these studies was a solid demonstration that tropotactic concentration-sensing algorithms could not explain the plume-tracking behavior in lobsters…So we are forced to consider other biologically feasible algorithms to find a reasonable explanation…Thus RoboLobster revealed to us something about the lobster's world that we had previously only suspected: the need to switch tracking strategies between different regions of the plume* [Grass02].

1.3 Academic Research activity

Using natural biology as a source of inspiration for solving sensing problems requires a solid understanding of biology. Although there is a long history of our understanding how biological sensory systems perform certain tasks, there is still very much that is not yet understood. Biological research institutions continually reveal deeper knowledge of the structure and function of sensory systems, which gives engineering problem-solvers more to consider. Models and algorithms are developed to match measured data, such as the Hassenstein-Reichardt Elementary Motion Detection (HR-EMD) model [Hass56], DeValois spatial vision models [DeV88], and others more focused on a specific sensory system, such as Frank Werblin's efforts to simulate primate vision processing in the retina [Werb91], John Douglas' and Nicholas Strausfeld's work to map the neural circuitry of the fly [Doug00], and many others.

Sometimes biology is deliberately considered for inspiration for new ideas. One example is the funding provided during the 1980's by the Office of Naval Research (ONR) and the Defense Advanced Research Project Agency (DARPA) to pursue novel military sensor designs. One of the products is a collection of biologically-inspired sensory system design concepts implemented in VLSI technology. Several of these designs developed at California Institute of Technology (CalTech) are detailed in *Analog VLSI and Neural System* [Mead89]. One of these designs, the 'silicon retina', was expanded by the Air Force Research Laboratory (AFRL) for military seeker applications by integrating with an array of infrared sensors [Mass93]. Some graduates from Mead's lab began their own bio-inspired labs at institutions such as Georgia Institute of Technology, University of Florida, Massachusetts Institute of Technology, etc. while other graduates started their own

companies building bio-inspired components or researching follow-on design concepts.

Much more work in bio-inspired sensing can be found in technical journals and conferences such as the *IEEE International Conference on Robotics and Biomimetics*. This conference alone has more than 500 papers and has been an annual conference since 2012. A common application for this conference is robotic fish, which has its inspiration in the design of fish for underwater maneuverability. Other popular topics include deep-learning neural networks, actuators, flocking (or swarming), and biomimetic materials. These topics are popular in other bio-inspired conferences, journals magazines, etc. Although the original neural network is bio-inspired, many subsequent efforts deviate from biology (not to mention very little is known about how real neural networks work). Any research using a neural network or adding something to a robotic fish or other originally bio-inspired concept could arguably be labeled 'bio-inspired', which complicates isolating truly new bio-inspired contributions.

In addition to the technology applications we have the more biology-focused efforts, where biologists are attempting to derive models that adequately reflect measured data. Example journals include *Vision Research* and *Biological Cybernetics*. A drawback for engineers is the biology-intensive language necessary to convey their models, as well as efforts typically are very focused on a very specific part of one species' neural circuitry, such as the mechanism for turning in the salamander [Liu20]. Therefore due to the quantity of technical efforts and the wide diversity of disciplines that consider this general topic area it is quite challenging to encompass all significant efforts in any of the basic modalities (vision, olfaction, gustation, tactile, audition) of bio-inspired sensory design.

Chapter 1 Questions

1. What do the terms *biomimetics, biometrics, biomedical, bionics,* and *bioinformatics* mean?
2. What are some of the motivations for studying biomimetics?
3. What terms do the various academic disciplines use to describe system structure and system function?
4. What are some of the characteristics of both artificial and natural neural networks?
5. There is a wealth of knowledge available concerning biological processing at the molecular and neuronal levels as well as a wealth of knowledge concerning human behavior. What is missing?
6. Why is it a good assumption that there will continue to be significant spending on biological research?

Chapter 1 References

[DeV88] DeValois, R. and DeValois, K, *Spatial Vision*, Oxford University Press, New York, 1988.

[Doug00] Douglass, J. and Strausfeld, N., "Optic flow representation in the optic lobes of Diptera: modeling the role of T5 directional tuning properties", *Journal of Comparative Physiology*, Vol. 186, No. 9, Sept. 2000.

[Grass02] Grasso, F., "Flow and Chemo-Sense for Robot and Lobster Guidance in Tracking Chemical Sources in Turbulence", Chapter 27 in Ayers, J., Davis, J. L., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, Cambridge, Massachusetts, 2002.

[Hass56] Hassenstein, W. R., "Systemtheoretische Analyse der Zeit-, Reihenfolgen- und Vorzeichenauswertung bei der Bewegungsperzeption des Rüsselkäfers Chlorophanus", *Zeitschrift für Naturforschung,* Vol. 11, pp. 513–524, 1956.

[Liu20] Liu, Q., et al., "Modeling of the neural mechanism underlying the terrestrial turning of the salamander", *Biological Cybernetics*, Vol. 114, No. 3, June 2020.

[Mass93] Massie, M., et al., "Neuromorphic infrared focal plane performs on-plane local contrast enhancement, spatial and temporal filtering", *Proceedings of the SPIE*, Vol. 1961, No. 17, 1993.

[Mead89] Mead, C., *Analog VLSI and Neural Systems*, Addison-Wesley, ISBN: 0-201-05992-4, 1989.

[Srini02] Srinivasan, M. V., "Visual Flight Control and Navigation in Honeybees: Applications to Robotics," in Ayers, J., Davis, J. L., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, Cambridge, Massachusetts, 2002.

[Werb91] Werblin, F. and Teeters, J., "Real-time simulation of the retina allowing visualization of each processing stage", *Proceedings of the SPIE*, Vol. 1472, 1991.

Chapter 2: GENERAL CONCEPTS FROM ENGINEERING AND BIOLOGY

Some common groundwork is necessary before investigating natural and biomimetic sensory systems and signal processing. A review of the salient aspects of linear systems (Section 2.1) is covered first. This is followed by fundamentals of neuronal systems (Section 2.2), neuronal processing (Section 2.3), an electric circuit model of a neuron (Section 2.4), and basic neuronal motion detection models (Section 2.5). The following free texts are recommended for more thorough treatment of linear systems theory and image processing:

Ulaby, F. and Yagle, A, *Signals and Systems: Theory and Applications*, Michigan Publishing, ISBN 978–1–60785–487–6, 2018. Available at ss2.eecs.umich.edu.

Yagle, A. and Ulaby, F., *Image Processing for Engineers*, Michigan Publishing, ISBN 978–1–60785–489–0, 2018. Available at ip.eecs.umich.edu.

2.1 Relevant Linear Systems Theory

This section summarizes the salient points of linear systems theory relevant to biomimetic sensory systems and signal processing. Many facets of natural vision processing can be modeled as spatial-temporal filters operating on input signals or image sequences. The two-dimensional spatial filters operate on each image frame, which is subsequently modified to account for the temporal history of previously filtered image frames. These operations are extensions of 1D discrete-time convolutions and other signal processing operations. The topics covered here include the motivation for LTI system modeling, continuous-time convolution and impulse response, $\delta(t)$, discrete-time convolution and unit pulse function, $\delta[n]$, 2D discrete-time convolution, and the Fourier Series and Fourier Transform.

2.1.1 Motivation for linear time-invariant (LTI) system modeling

Biological systems are naturally non-linear and time-varying. However, there is much practicality in approximating portions of the system as piece-wise linear over a nominal range of input values and time-invariant for relatively short periods of time. Such models many times can be "close enough" to be very useful. To simplify signal processing, we desire to use models of biological information transforms that are *linear* and *time-invariant*, or LTI transforms.

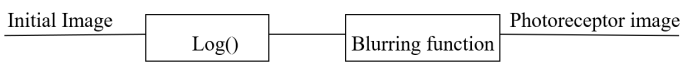
A system is *linear* if the law of superposition applies. In electrical engineering, superposition implies mathematical *homogeneity* and *additivity*. If the system output is $y_1 = f(x_1)$ for an input x_1 and $y_2 = f(x_2)$ for an input x_2, then the system is homogeneous if $f(\alpha x_1) = \alpha y_1$, additive if

$$f(x_1 + x_2) = f(x_1) + f(x_2) = y_1 + y_2$$

and therefore, linear if $f(\alpha x_1 + \beta x_2) = f(\alpha x_1) + f(\beta x_2) = \alpha y_1 + \beta y_2$

A system is *time-invariant* if for the same response is given for the same set of inputs, regardless of when the inputs are presented. That is, for $y(t) = f(x(t))$, then $f(x(t - t_0)) = y(t - t_0)$ for a constant time interval t_0.

Many natural (biological) signal processing functions can be modeled as a sequence of LTI subsystems we refer to as *filters*. The signal is sent through a filter and looks different at the output. For example, a filter simulating the layer of photoreceptors in mammalian vision systems may include a logarithmic conversion of light intensity followed by a blurring effect due to interactions with nearby photoreceptors. This model would include a logarithmic filter followed by a Gaussian blurring filter:



If $h(t)$ is the *impulse response* to the system, then, for a continuous-time input signal $f(t)$

$$h(t) = f(t) * \delta(t) = \int_0^t f(\tau)\delta(t-\tau)d\tau$$

and for a discrete-time input sequence $f[n]$

$$h[n] = f[n] * \delta[n] = \sum_{k=-\infty}^{\infty} f[k]\delta[n-k]$$

The symbol * is used to denote convolution, and δ denotes the Direct delta function, both of which are discussed later. For the discrete-time case, time-invariance implies that if *h[n]* is the response to *δ[n]*, then *h[n-k]* is the response to *δ[n-k]*. The impulse response (*h[n]* for discrete-time systems and *h(t)* for continuous-time systems) of a linear time-invariant (LTI) system completely characterizes that system. The output, *y[n]*, of an LTI system is the sum of individual impulse responses weighted by the current input value, *x[n]*:

$$y[n] = \sum_{k=-\infty}^{\infty} x[k]h[n-k]$$

For a continuous-time signal, *s(t)*, the *Fourier Transform* of *s(t)*, represented as *S(ω)*, shows the frequency content of *s(t)* as a function of radian frequency ω. The *Inverse Fourier Transform* of *S(ω)* gives back the original time-domain representation. The two functions, both representing the same signal from different domain perspectives form a Fourier Transform pair, denoted as: *s(t) <==> S(ω)*.

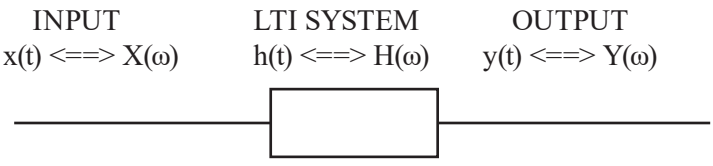
The response of a system to an impulse is called the *impulse response* and is denoted as *h(t)*. The frequency representation of the impulse response, *H(ω)*, is called the *frequency response*. Given an input signal to a known LTI system, the output can be determined:

in the time domain as the convolution of input and LTI impulse response, or

in the frequency domain as the product of the input spectrum and the LTI frequency response.

To put it another way, given an input, *x(t)*, whose spectrum is *X(ω)*, to an LTI systems whose impulse response is *h(t)* (whose spectrum is *H(ω)*, the frequency response) then the output, *y(t)*, and its spectrum, *Y(ω)*, is given as

Time domain: $y(t) = x(t) * h(t)$
Frequency domain: $Y(\omega) = X(\omega)\,H(\omega)$



<==> Fourier Transform pair
h(t) – Impulse Response
H(ω) – Frequency Response

y(t) = x(t) * h(t)
Y(ω) = X(ω) H(ω)

2.1.2 Continuous-time convolution

For continuous functions, *f(t)* and *g(t)*, the *convolution* is defined by

$$f(t) * g(t) = \int_0^t f(\tau)g(t-\tau)d\tau$$

and has these properties:

$f * g = g * f$	commutative property
$f * (g_1 + g_2) = f * g_1 + f * g_2$	distributive property
$(f * g) * v = f * (g * v)$	associative property

2.1.3 Continuous-time unit impulse function

The *unit impulse function* or *Dirac delta function* is defined as

$\delta(t) = \infty,\ t = 0$ $\quad = 0,\ t \neq 0$ $\qquad$ $\delta(t-a) = \infty,\ t = a$ $\quad = 0,\ t \neq a$

and

$$\int_{-\infty}^{\infty} \delta(t)dt = \int_{-\infty}^{\infty} \delta(t-a)dt = 1$$

since *δ(t-a)* has infinite strength at $t = a$, has zero duration at $t = a$, and has unity area.

A continuous-time signal *x(t)* can be replicated by the convolution with the unit impulse function:

$$x(t) = \int_{-\infty}^{\infty} x(\tau)\delta(t-\tau)d\tau$$

2.1.4 Discrete-time unit pulse function

A continuous-time signal can be discretized into a sequence, *x*[0], *x*[1], *x*[2] ..., by collecting the values of this integral at the specific times t_0, t_0+T, t_0+2T, t_0+3T, etc. so that

$x[0] = x(t)*\delta(t_0) = x(t_0)$
$x[1] = x(t)*\delta(t_0+T) = x(t_0+T)$
$x[2] = x(t)*\delta(t_0+2T) = x(t_0+2T)$
: : :
$x[n] = x(t)*\delta(t_0+nT) = x(t_0+nT)$

Brackets, [ ], are used to denote discrete-time sequences while parentheses, ( ), are used to denote continuous-time functions. In this example, the sequence *x*[0], *x*[1], ...*x*[n]...

is a discrete-time representation of the continuous-time function $x(t)$.

The discrete-time version of the unit impulse sequence is called the *unit sample sequence* or the *unit pulse function*. The unit pulse function is defined as

$$\delta[n] = \begin{cases} 1, & n = 0, \\ 0, & n \neq 0 \end{cases}$$

To properly interpret between continuous-time functions and discrete-time sequences, it should be noted that the unit pulse function must have *unity* area. To maintain this definition for a pulse of duration T, then the amplitude must be defined as $1/T$. The definition above presumes $T = 1$.

A discrete-time signal $x[n]$ can be replicated by the convolution with the unit pulse function:

$$x[n] = x[n] * \delta[n] = \sum_{k=-\infty}^{\infty} x[k]\delta[n-k]$$

2.1.5 One-dimensional discrete-time convolution

The discrete-time convolution operation is defined as

$$f[n] * g[n] = \sum_{k=-\infty}^{\infty} f[k]g[n-k]$$

Equation 2.1-1

Example 2.1-1

Use Equation 2.1-1 to calculate $y[n] = f[n] * g[n]$ where $f[n] = \{1\ \ 2\ \ 3\}$ and $g[n] = \{1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\}$.

Solution:

$f[n] = \{1\ \ 2\ \ 3\}$ implies $f[0] = 1, f[1] = 2, f[2] = 3$ and $g[n] = \{1\ \ 1\ \ 1\ \ 1\ \ 1\ \ 1\}$ implies $g[0] = g[1] = g[2] = g[3] = g[4] = g[5] = 1$

For n = 0, Equation 2.1-1 gives the summation $y[0] = \cdots f[-1]g[1] + f[0]g[0] + f[1]g[-1] + \ldots$

but all values of $f[n]$ and $g[n]$ are zero for $n < 0$, so $y[0] = f[0]g[0] = 1$.

$$y[0] = \sum_{k=0}^{5} f[k]g[0-k] = f[0]g[0] = 1$$

$$y[1] = \sum_{k=0}^{5} f[k]g[1-k] = f[0]g[1] + f[1]g[0] = 1 + 2 = 3$$

$$y[2] = \sum_{k=0}^{5} f[k]g[2-k] = f[0]g[2] + f[1]g[1] + f[2]g[0] = 1 + 2 + 3 = 6$$

$$y[3] = \sum_{k=0}^{5} f[k]g[3-k] = f[0]g[3] + f[1]g[2] + f[2]g[1] = 1 + 2 + 3 = 6$$

$$y[4] = \sum_{k=0}^{5} f[k]g[4-k] = f[0]g[4] + f[1]g[3] + f[2]g[2] = 1 + 2 + 3 = 6$$

$$y[5] = \sum_{k=0}^{5} f[k]g[5-k] = f[0]g[5] + f[1]g[4] + f[2]g[3] = 1 + 2 + 3 = 6$$

$$y[6] = \sum_{k=0}^{5} f[k]g[6-k] = f[1]g[5] + f[2]g[4] = 2 + 3 = 5$$

$$y[7] = \sum_{k=0}^{5} f[k]g[7-k] = f[2]g[5] = 3$$

Where products are omitted when they are zero for specific k values. The result is $y[n] = \{1\ 3\ 6\ 6\ 6\ 6\ 5\ 3\}$.

■

We could also visualize the answer graphically by reversing the sequence $g[n]$ and placing it below $f[n]$ and offsetting by the value of n in Equation 2.1–1. The first three values determined using this graphical approach are shown here:

***n* =0:**

k:	−5	−4	−3	−2	−1	0	1	2	3	4	5
f[k]:	0	0	0	0	0	1	2	3	0	0	0
g[n-k]:	1	1	1	1	1	1	0	0	0	0	0

$$y[0] = \sum_{k=0}^{5} f[k]g[0-k] = f[0]g[0] = 1$$

***n* =1:**

k:	−5	−4	−3	−2	−1	0	1	2	3	4	5
f[k]:	0	0	0	0	0	1	2	3	0	0	0
g[n-k]:	0	1	1	1	1	1	1	0	0	0	0

$$y[1] = \sum_{k=0}^{5} f[k]g[1-k] = f[0]g[1] + f[1]g[0] = 1 + 2 = 3$$

***n* =2:**

k:	−5	−4	−3	−2	−1	0	1	2	3	4	5
f[k]:	0	0	0	0	0	1	2	3	0	0	0
g[n-k]:	0	0	1	1	1	1	1	1	0	0	0

$$y[2] = \sum_{k=0}^{5} f[k]g[2-k] = f[0]g[2] + f[1]g[1] + f[2]g[0] = 1 + 2 + 3 = 6$$

And so on.

Example 2.1–2

Solve the same convolution (Example 2.1–1) using Equation 2.1–1 but reverse the roles of $f[n]$ and $g[n]$.

Solution:

Due to the commutative property the result should be the same. Using Equation 2.1-1, the order could also be reversed, as $y = g * f$, giving:

$$y[0] = \sum_{k=0}^{5} g[k]f[0-k] = g[0]f[0] = 1$$

$$y[1] = \sum_{k=0}^{5} g[k]f[1-k] = g[0]f[1] + g[1]f[0] = 2 + 1 = 3$$

$$y[2] = \sum_{k=0}^{5} g[k]f[2-k] = g[0]f[2] + g[1]f[1] + g[2]f[0] = 3 + 2 + 1 = 6$$

$$y[3] = \sum_{k=0}^{5} g[k]f[3-k] = g[1]f[2] + g[2]f[1] + g[3]f[0] = 3 + 2 + 1 = 6$$

$$y[4] = \sum_{k=0}^{5} g[k]f[4-k] = g[2]f[2] + g[3]f[1] + g[4]f[0] = 3 + 2 + 1 = 6$$

$$y[5] = \sum_{k=0}^{5} g[k]f[5-k] = g[3]f[2] + g[4]f[1] + g[5]f[0] = 3 + 2 + 1 = 6$$

$$y[6] = \sum_{k=0}^{5} g[k]f[6-k] = g[4]f[2] + g[5]f[1] = 3 + 2 = 5$$

$$y[7] = \sum_{k=0}^{5} g[k]f[7-k] = g[5]f[2] + g[5]f[1] = 3$$

so that once again $y[n] = \{1\ \ 3\ \ 6\ \ 6\ \ 6\ \ 6\ \ 5\ \ 3\}$.

■

As before we could also visualize the answer graphically, but this time by reversing the sequence $f[n]$ and placing it below $g[n]$ and offsetting by the value of n in Equation 2.1-1. The first three values determined using this graphical approach are shown here:

***n* =0:**

k:	−2	−1	0	1	2	3	4	5	6
g[k]:	0	0	1	1	1	1	1	1	0
f[n-k]:	3	2	1	0	0	0	0	0	0

$$y[0] = \sum_{k=0}^{5} g[k]f[0-k] = g[0]f[0] = 1$$

***n* =1:**

k:	−2	−1	0	1	2	3	4	5	6
g[k]:	0	0	1	1	1	1	1	1	0
f[n-k]:	0	3	2	1	0	0	0	0	0

$$y[1] = \sum_{k=0}^{5} g[k]f[1-k] = g[0]f[1] + g[1]f[0] = 2 + 1 = 3$$

***n* =2:**

k:	−2	−1	0	1	2	3	4	5	6
g[k]:	0	0	1	1	1	1	1	1	0
f[n-k]:	0	0	3	2	1	0	0	0	0

$$y[2] = \sum_{k=0}^{5} g[k]f[2-k] = g[0]f[2] + g[1]f[1] + g[2]f[0] = 3 + 2 + 1 = 6$$

And so on.

Convolutions can be computed quickly in *MatLab*, which stands for "matrix laboratory", a software product of Mathworks, Inc. [MatLab]. The sequence $y = f * g$ {implying $y[n] = f[n]*g[n]$} as defined in Example 2.1–1 is computed in Matlab as

```
>> f = [1  2  3];
>> g = [1  1  1  1  1  1];
>> y = conv(f,g)
y =   1    3    6    6    6    6    5    3
```

Note that the resulting sequence is longer that both input sequences; this is a natural consequence of the convolution operation. Also, MatLab does not use a zeroth element; the first element of y is referenced in MatLab as $y[1]$, which

is the same value as $y[0]$ in Equation 2.1–1. Given these considerations the simulation confirms the hand-written results in Example 2.1–1.

Exercise 2.1–1

Use Equation 2.1-1 to calculate $y[n] = x[n] * h[n]$ for $x[n] = \{1\ \ 2\ \ 2\}$ and $h[n] = \{1\ \ 3\}$. Check your answer using *graphical* convolution.

Answer: $y[n] = \{1\ \ 5\ \ 8\ \ 6\}$

■

Exercise 2.1–2

Use Equation 2.1-1 to calculate for the given sequences. Check your answer using graphical convolution.

a) $x[n] = \{1\ \ 1\ \ 1\ \ 1\ \ 1\}$ and $h[n] = \{0.25\ \ 0.5\ \ 0.25\}$.
b) $x[n] = \{1\ \ 1\ \ 1\ \ 1\ \ 1\}$ and $h[n] = \{0.25\ \ -0.5\ \ 0.25\}$.

Answers:

a) $y[n] = \{0.25\ \ 0.75\ \ 1.0\ \ 1.0\ \ 1.0\ \ 0.75\ \ 0.25\}$
b) $y[n] = \{0.25\ \ -0.25\ \ 0\ \ 0\ \ 0\ \ -0.25\ \ 0.25\}$

■

2.1.6 Two-dimensional discrete-time convolution

In signal processing applications, one sequence may represent a filter and another a given input signal. Convolving the two signals would give an output that is a filtered representation of the given input signal. Image processing is two-dimensional (2D) signal processing, where one 2D signal (image) represents a filter and the other the input image. For biomimetic applications, the filter could represent a model of a natural phenomenon as it affects the raw imagery. Assuming 2D variables *f(x,y)* and *g(x,y)*, the 2D convolution operation is given as:

$$f[x,y]*g[x,y] = \sum_{m=0}^{M-1}\sum_{n=0}^{N-1} f[m,n]g[x-m,y-n]$$

Example 2.1–3

Using MatLab computer the 2D convolution $y = f * g$ given the 2D variables f and g defined here:

f = 1 1 / 1 1 g = 2 2 2 / 2 2 2 / 2 2 2

Solution:

```
>> f = [1 1; 1 1];
>> g = [2 2 2; 2 2 2; 2 2 2];
>> y = conv2(f,g)
```

y =

2	4	4	2
4	8	8	4
4	8	8	4
2	4	4	2

■

Notice that for both 1D and 2D convolutions the result tends to grow. Sometimes it is necessary to 'crop' out the internal pixels so that the filtered version is the same size as the original. For example, defining *f* below as a 5x5 2D unit pulse and performing the convolution with the 3x3 2D variable *g* results in a 7x7 result:

```
>> f = zeros(5);
>> f(3,3) = 1;
```

f =

0	0	0	0	0
0	0	0	0	0
0	0	1	0	0
0	0	0	0	0
0	0	0	0	0

```
>> g = [1 2 3; 4 5 6; 7 8 9];
```

g =

1	2	3
4	5	6
7	8	9

```
>> y=conv2(f,g);
```

y =

0	0	0	0	0	0	0
0	0	0	0	0	0	0
0	0	1	2	3	0	0
0	0	4	5	6	0	0
0	0	7	8	9	0	0
0	0	0	0	0	0	0
0	0	0	0	0	0	0

To visualize the computation, flip g both vertically and horizontally: $g^* =$

$$\begin{matrix} 9 & 8 & 7 \\ 6 & 5 & 4 \\ 3 & 2 & 1 \end{matrix}$$

The values for y are found by placing g^* over each element of f, and then performing a dot product; that is, multiply element-for-element, and then add the products (sum of products). One way to crop out the middle 5x5 is to manually redefine y:

```
>> y = y(2:6,2:6);
```

This command redefines y to be row 2 through 6 and column 2 through 6. The result is cropping the internal 5x5 image of the previous convolution so that the filtered version is now the same size as the original image, *f*.

```
y =     0     0     0     0     0
        0     1     2     3     0
        0     4     5     6     0
        0     7     8     9     0
        0     0     0     0     0
```

A better alternative for cropping out the central pixels is to use the *same* attribute in the *conv2* Matlab function when determining y:

```
>> y=conv2(f,g,'same');   y =   0     0     0     0     0
                               0     1     2     3     0
                               0     4     5     6     0
                               0     7     8     9     0
                               0     0     0     0     0
```

Exercise 2.1–3

Given the 2D filter f and image x, give the output filtered image $y = f * x$.

$$f = \begin{matrix} -2 & 0 & 2 \\ -1 & 0 & 1 \\ 0 & 1 & 0 \end{matrix} \qquad x = \begin{matrix} 0 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{matrix}$$

Give both the cropped and uncropped representations of the output.

Answers:

Uncropped: y =

$$\begin{matrix} 0 & 0 & 0 & 0 & 0 \\ 0 & -2 & -2 & 2 & 2 \\ 0 & -1 & -1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{matrix}$$

Cropped: y =

$$\begin{matrix} -2 & -2 & 2 \\ -1 & -1 & 1 \\ 0 & 1 & 1 \end{matrix}$$

2.2 Neuronal Sensory Systems and Concepts

One of the difficulties in exploiting natural neuronal structure and function for engineering applications is the lack of understanding between the microscopic physiology and macroscopic behavior. *Neuroscience* is the study of neurons at the microscopic level, while *psychophysics* is the study of correlations between specific physical stimuli and the sensations that result. Neuroscience is somewhat focused on the chemistry of the neuron and psychophysics is more focused on the macroscopic behavior of the complete organism. The story of how and why neurons are connected the way they are continues to unfold and will continue for centuries to come.

2.2.1 Massive Interconnections

Neurons are highly interconnected to form the information channels in sensory systems. For example, the human brain contains about 10^{10} to 10^{12} neurons, each making up to 10^3 to 10^4 connections to other neurons. Small groups of neurons are called *ganglia*, which typically controls specific behaviors of an animal. Many invertebrates have large neurons and small ganglia, which allows researchers an opportunity to investigate neuronal signaling and primitive neuronal networks.

2.2.2 Hebbian Learning

Natural neuronal connections are often strengthened with continued use, known as *Hebbian Learning*. The result is an adaptation of the network to the most frequent signal sequences from external stimuli. *Artificial neural networks* (ANN) are now commonly used to solve computational problems when direct analytical methods are difficult or impossible. These networks are inspired by the natural neuronal paradigm, and many have taken on variations that diverge from these original examples. This is quite acceptable since the typical goal is to solve some engineering or computational problem, not necessarily to mimic the natural paradigm.

2.2.3 Physical Types of Natural Sensors

There are different ways to categorize the natural neuronal systems designs that are available for engineering exploitation. The method chosen here is based on the physics of the stimulus. The ones given the most attention are the ones most commonly found in nature:

- Photo-sensory systems, stimulated by photons
  -- vision systems in vertebrates and invertebrates
- Mechano-sensory systems, stimulated by physical motion in the environment
  -- touch systems in vertebrates and invertebrates
  -- auditory systems in vertebrates and invertebrates
  -- kinesthesia, which is knowing the relative positions of body parts
- Chemo-sensory systems, stimulated by changes in chemical content of stimuli
  -- olfactory systems providing the sense of smell
  -- gustation systems providing the sense of taste

Other physical senses occasionally found in biology include those sensitive to heat, infra-red radiation, polarized light, electric fields, and magnetic fields.

There are three basic types of stimulus reception in biological sensory systems:

- *Exteroception* is the receiving of signals from outside the organism, such as photons of light for the vision system, sound waves for the auditory system, and chemical traces for the olfactory system. Sensory systems in this group are the subject of this text and most of the bio-inspired sensory system research that has been done.
- *Proprioception* is the receiving of signals that relate position of body segments to one another and the position of the body in space, which involves kinesthesia mentioned earlier.
- *Interoception* is the receiving of signals from conditions inside the organism, such as blood glucose level and blood pressure level.

There are three basic maps of sensory receptive fields to portions of the brain:

- *Somatotopic Map* is a map of the body surface in the somatosensory cortex.
- *Retinotopic Map* is a map of the visual field (as focused onto the retina) in the primary visual cortex in the occipital lobe of the brain.
- *Tonotopic Map* is a map of the basilar membrane in the primary auditory cortex in the temporal lobe of the brain.

The amount of brain surface area dedicated to various regions of reception varies dramatically, as certain reception areas are more important and require more dedicated processing. For example, the allocation in the brain on the somatotopic map for the sensation of touch in the index finger is much larger than the same relative skin surface area of the back. Another interesting point is the nearest-neighbor receptor mapping is generally preserved in the cortex. That is, adjacent receptors in the peripheral sensory system tend to stimulate adjacent neurons in the cortex.

2.3 Fundamentals of neuronal processing

This section reviews basic neuronal topics that are prevalent in biological sensory systems such as the vision system, the auditory system, and the olfactory system. Although biological neurons are much slower than modern transistor electronics, the fundamental principles of neuronal processing exploit natural logarithmic behavior of charge distributions and transport. The *pn* junction exhibits this natural relationship, but we tend to take pairs of *pn* junctions (transistors) and create a binary switch (digital bit). If the exponential *v-i* relationship of a *pn* junction could be used at today's computer clock speeds, there could be many orders of magnitude improvement in computational performance and power consumption. In addition, there is still much to be learned about the interconnection strategies found in natural neuronal networks.

2.3.1 Adaptation and Development

Although biological systems can be studied as existing systems that solve processing problems, it should be noted that these systems are constantly developing and adapting. From conception to death every known biological system is continually maturing, never reaching an unchanging physical state. The neuronal system of a mature adult is

relatively stable, thus representing some level of neuronal optimization due to environmental adaptation.

Adaptation can be immediate or long term. An example of immediate adaptation is the response of the iris of the eye to light levels, controlling the amount of photonic flux entering the pupil. A short-term adaptation, called *habituation* and *sensitization*, is demonstrated in the marine snail, *Aplysia*. The gill is withdrawn beneath a mantle in defense when the siphon, attached to the mantle, is stimulated. The reflex magnitude is decreased as the siphon is artificially stimulated, resulting in the habituation, or desensitization, of the response to the experimental environment. The response can be subsequently sensitized by stimulating other parts of the body. Through training these reflex conditions can be made to last for days, indicating a primitive form of memory and learning [Dowl87].

In higher life forms these simple neuronal adaptations combine with massive interconnections to provide more complex adaptation concepts. For example, it has been demonstrated that detection of spatial harmonics and identification of complicated sinusoidal grating patterns depends on adaptation to the harmonics and harmonic patterns. The detection threshold increases after adaptation to the harmonic, and the pattern identification threshold increases after adaptation to the patterns [Vass95].

Training during one's lifetime is an example of long-term adaptation. Training results in neurons being connected or strengthened, which is called *coincidence learning* or *Hebb learning* [Hech90]. A more long-term adaptation is genetic coding, passing adaptation information from one generation to the next.

2.3.2 Sense Organs and Adaptation

The following *italicized* text concerning the sense organs in crabs is quoted from [Warner77] (non-italicized text is additional commentary):

> *"Sense organs function at the cellular level by converting the stimulus into a change in the electrical potential across the receptor cell membrane. This receptor potential, if sufficiently large, results in the initiation of nerve impulses (action potentials) which are transmitted along nerve to the CNS"* (Central Nervous System)*."*

In some cases, however, such as primate vision, there are additional layers of cells between the receptors and action-potential transmission axons. In these instances, graded preprocessing functions occur before the information is encoded into action potentials. But for the simpler sensory system designs, the action potential...

> *"...frequency is a measure of the strength of the stimulus...Each receptor cell is specialized to convert a particular type of stimulus (light, mechanical deformation, etc.) and each has a particular threshold below which the stimulus is insufficient to trigger nerve impulses. Maintained stimulation generally results in the threshold of a receptor being raised (i.e. the receptor becomes less sensitive). Thus, many receptor cells, described as rapidly adapting, respond with a short burst of impulses only at the initiation of stimulation. Others which respond over longer periods of maintained stimulation are referred to as slowly adapting or, in extreme cases, non-adapting. Single sense organs may be composed of several receptor cells each with a different rate of adaptation."*

To illustrate slow and rapid adaptation, consider two neurons whose threshold for action-potential generation is –55mV. As ions entered the neuron from input dendrites, the membrane potential would increase from a resting potential (about –70 mV) to the –55 mV threshold. An action potential spike would then cause positive ions to be discharged during the spike, resulting in the membrane potential returning to the resting potential. Then the process would start over, where incoming ions would build up the membrane potential to the threshold for the initiation of an action potential. For rapid adaptation, the threshold might rise significantly after each spike generation, and degrade rapidly back to –55 mV when the input stimulus is removed. For slow adaptation, the threshold might rise nominally after each spike generation, and degrade more slowly back to –55 mV once the stimulus is removed.

Figure 2.3.2–1 shows the results of a neuronal adaptation model that accounts for increases in action-potential threshold with firing activity as well as a return to a nominal resting threshold in the absence of input stimulus. The input rectangular waveform is amplified to show when the stimulus is on and when it is off. As the neuronal membrane potential increases with input stimulus, there is a constant leakage that tends to bring the potential back to its resting state (about –70mV). Similarly, there is a constant leakage in the threshold that tends to bring it back to its resting state (about –55 mV). As the neuron adapts to the stimulus, the threshold level is raised, which subsequently reduces the action potential (spike) frequency.

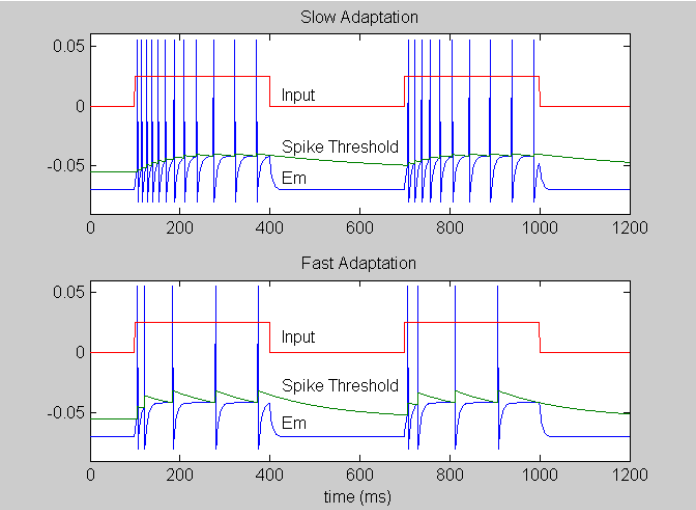


Figure 2.3.2–1.
Slow and Fast Neuronal Adaptation.

In this model, the speed of the model is implemented by the jump in threshold level after an action potential and the rate the threshold returns to resting state when stimulus stops. The input curve is amplified to better show when the input is present and when it is absent. Thr_adapt = 0.002 (slow) and 0.01 (fast) represents increase in threshold after action potential, and Thr_leak = .003 (slow) and 0.006 (fast) represents threshold return rate to –55 mV.

Initially, the spike frequency is high; as the neuron adapts, it is reduced, as seen in the figure. During slow adaptation, the threshold is not increased very much after each spike so that it takes longer for the frequency to reach a steady-state low frequency. During fast adaptation, the threshold is increased significantly after each spike so that the neuron quickly reaches a steady-state low frequency. An alternate mechanism for rapid adaptation is for the neuron to return to a potential higher than the normal resting potential after an action potential. Instead of returning to –70 mV, it may only return to, say, –60 mV. Then the neuron has a much smaller potential increase to obtain before reaching the next threshold level for action potential firing.

2.3.3 Ionic Balance of Drift and Diffusion

A fundamental principle of semiconductor physics is the balance of charge carrier drift and diffusion currents in the depletion region of a *pn* junction. Both materials are characterized by a high concentration of mobile carriers (holes for *p*-type, electrons for *n*-type) within an electrostatically neutral volume. When brought together, electrons in the *n*-type material diffuse from a higher concentration area, leaving charge-depleted positively charged lattice ions and combine in the p-type material to form a negatively charged lattice ion there. As this process continues, an electric field from the positive lattice (*n*-type) to the negative lattice (*p*-type) is growing in strength. This field causes carriers to retreat, thus developing a drift current opposing diffusion current. Equilibrium is reached when the diffusion current equals the drift current [Horen96].

Biological cells are also held in equilibrium in part by a balance of drift and diffusion currents. However, the charge carriers are primarily potassium, K^+, sodium, Na^+, and chlorine Cl^- ions, which are not as mobile as holes and electrons. The electronic currents are defined in terms of diffusion and mobility device constants, functions primarily of doping and geometry [Horen96]. Biological currents are functions of cell membrane permeabilities, which are functions of time, membrane potential, and ionic concentrations [MacG91].

In the case of potassium and sodium, there is a separate organic mechanism called the *Na-K Pump* that keeps ionic concentrations stable inside the neuron. The Na-K Pump uses metabolic energy supplied by the stored biological energy in the organism. For chlorine, however, the interior and exterior concentrations are maintained close together, balanced by drift and diffusion.

2.3.4 Nernst and Goldman Equations

An electric potential, or voltage, is established across a membrane when there is an unequal concentration of ions in the two regions. The Nernst equation (1908) showing this relationship is given by

$$E = \frac{RT}{nF} ln \frac{[C]_o}{[C]_i} = \frac{kT}{zq} ln \frac{[C]_o}{[C]_i}$$

Nernst Equation

$$\simeq (26mV) ln \frac{[C]_o}{[C]_i}, \; for \; K^+, \; Na^+$$

$$\simeq -(26mV) ln \frac{[C]_o}{[C]_i} \simeq (26mV) ln \frac{[C]_i}{[C]_o}, \; for \; Cl^-$$

where R is the universal gas constant, T is the absolute temperature, F is the Faraday constant (electrical charge per gram equivalent ion), n is the charge on the ion and $[C]_o$ and $[C]_i$ are the ion concentrations outside and inside the cell. Note that $\frac{[RT]}{[F]} = \frac{[kT]}{[q]} \approx 26mV$ at room temperature. The latter term is used frequently in electronic circuits.

Bernstein (1912) presented the significance to neuroelectric signaling of ionic fluxes about neuronal membranes. Building on this concept, the Nernst equation, and fundamental physics of ionic media contributed by Planck and Einstein, Goldman (1958) contributed the primary model for resting potential in neurons as

$$E = \frac{kT}{q} ln\left(\frac{P_{K+}[K^+]_o + P_{Na+}[Na^+]_o + P_{Cl-}[Cl^-]_i}{p_{K+}[K^+]_i + P_{Na+}[Na^+]_i + P_{Cl-}[Cl^-]_o} \right)$$

Goldman Equation

where T is absolute temperature, k is Boltzmann's constant, q is unitary electric charge, P's are permeabilities, $[\ ]_o$'s are concentrations outside the cell, and $[\ ]_i$'s are concentrations inside the cell. As long as the permeabilities are constant, the Goldman equation gives good steady-state results [MACG91].

Example 2.3.4–1

Find the Nernst potential between the interior and exterior of a neuron due to each ion if their concentrations are as follows:

	Inside	Outside
K^+	360 mM	20 mM
Na^+	45 mM	450 mM
Cl^-	50 mM	600 mM

(M stands for *mole*, a unit of concentration)

Solution:

Potential due to concentration differences alone are solved by the Nernst Equation, keeping in mind the ratio of outer-to-inner ion concentration are reversed for negatively-charged ions:

$$E_{K+} = (26mV)\ln\frac{[C]_o}{[C]_i} = 0.026\ln\frac{20}{360} = -75.1m$$

$$E_{Na+} = (26mV)\ln\frac{[C]_o}{[C]_i} = 0.026\ln\frac{450}{45} = 59.9m$$

$$E_{Cl-} = (26mV)\ln\frac{[C]_i}{[C]_o} = 0.026\ln\frac{50}{600} = -64.6m$$

■

Example 2.3.4–2

For the given ion concentration determine the membrane potential assuming the following ratio of ion permeabilities: $P_{K+} : P_{Na+} : P_{Cl-} = 1.0 : 0.5 : 0.2$

	Inside	Outside
K^+	320 mM	25 mM
Na^+	40 mM	420 mM
Cl^-	60 mM	540 mM

Solution:

The relative permeabilities given means that $P_{Na+} = 0.5P_{K+}$ and $P_{Cl-} = 0.2P_{K+}$; using the Goldman Equation

$$E_m = (0.026)\ln\left(\frac{P_{K+}[K^+]_o + P_{Na+}[Na^+]_o + P_{Cl-}[Cl^-]_i}{p_{K+}[K^+]_i + P_{Na+}[Na^+]_i + P_{Cl-}[Cl^-]_o} \right)$$

$$E_m = (0.026)\ln\left(\frac{P_{K+}[K^+]_o + 0.5P_{Na+}[Na^+]_o + 0.2P_{Cl-}[Cl^-]_i}{p_{K+}[K^+]_i + 0.5P_{Na+}[Na^+]_i + 0.2P_{Cl-}[Cl^-]_o} \right)$$

$$E_m = (0.026)\ln\left(\frac{[K^+]_o + 0.5[Na^+]_o + 0.2[Cl^-]_i}{[K^+]_i + 0.5[Na^+]_i + 0.2[Cl^-]_o} \right)$$

$$E_m = (0.026)\ln\left(\frac{25 + 0.5(420) + 0.2(60)}{320 + 0.5(40) + 0.2(540)} \right) = (0.026)\ln\frac{247}{448} = -15.48m$$

$$E_m = (0.026)$$

■

Example 2.3.4–3

Given the permeability ration $P_{K+} : P_{Cl-} = 1.0 : 0.5$ and the ion concentrations

	Inside	Outside
K^+	400 mM	20 mM
Na^+	50 mM	440 mM
Cl^-	52 mM	560 m

determine the relative permeability k of P_{Na+} such that $P_{K+} : P_{Na+} : P_{Cl-} = 1.0 : k : 0.5$ and the resting membrane potential is +50 mV.

Solution:

The relative permeabilities given means that $P_{Na+} = kP_{K+}$ and $P_{Cl-} = 0.5P_{K+}$; using the Goldman Equation

$$0.05 = (0.026)\ln\left(\frac{P_{K+}[K^+]_o + P_{Na+}[Na^+]_o + P_{Cl-}[Cl^-]_i}{p_{K+}[K^+]_i + P_{Na+}[Na^+]_i + P_{Cl-}[Cl^-]_o}\right)$$

$$0.05 = (0.026)\ln\left(\frac{P_{K+}[K^+]_o + k\,P_{Na+}[Na^+]_o + 0.5P_{K+}[Cl^-]_i}{p_{K+}[K^+]_i + k\,P_{Na+}[Na^+]_i + 0.5P_{K+}[Cl^-]_o}\right)$$

$$0.05 = (0.026)\ln\left(\frac{[K^+]_o + k[Na^+]_o + 0.5[Cl^-]_i}{[K^+]_i + k[Na^+]_i + 0.5[Cl^-]_o}\right)$$

$$0.05 = (0.026)\ln\left(\frac{20 + k(440) + 0.5(52)}{400 + k(50) + 0.5(560)}\right)$$

$$1.923 = \ln\left(\frac{440k + 46}{50k + 680}\right)$$

$$e^{1.923} = 6.84 = \frac{440k + 46}{50k + 680}$$

$$6.84(50k + 680) = 440k + 46 \Longrightarrow k = 47$$

■

Exercise 2.3.4–1

Find the Nernst potential for each ion using the concentrations given in Examples 2.3.4-2 and 2.3.4-3

Answers:

Example 2.3.4-2: E_K= -66.3 mV, E_{Na}= 61.1 mV, E_{Cl} = -57.1 mV
Example 2.3.4-3: E_K= -77.9 mV, E_{Na}= 56.5 mV, E_{Cl} = -61.8 mV

■

Exercise 2.3.4–2

a) Repeat Example 2.3.4-2 with permeability ratio $P_{K+} : P_{Na+} : P_{Cl-} = 1.0 : 5.0 : 0.2$

b) Repeat Example 2.3.4-2 with permeability ratio $P_{K+} : P_{Na+} : P_{Cl-} = 1.0 : 0.05 : 0.2$

Answers:

a) 31.8 mV

b) -52.1 mV

■

Exercise 2.3.4–3

Repeat Example 2.3.4-3 so that the resting membrane potential is -50 mV.

Answer: $k = 0.123$

■

2.3.5 The Action Potential

In general, neuronal processing within localized areas occurs as *graded* processing, while information transmission over reasonable lengths occurs as a frequency of asynchronous pulses, called *spikes*. In graded processing, the potential of the cell and its output rise slowly as the input signal levels strengthen. When input signals cease, then the cell and its output slowly return to the resting membrane potential. In spike-train processing, the inputs to the cell cause the interior to rise in potential until a certain threshold is reached. Typically, the resting potential is about –70mV (with respect to the extracellular fluid) and the threshold for initiating a spike is about –55 mV. The profile of voltage per time during a spike is known as the *action potential*.

When signals are transmitted over long distances, the signal transmission process tends to be a series of action potentials whose occurrences increase (higher frequency) as the input signal increases. Hodgkin and Huxley (1952) presented the original set of equations that describe the generation of a single action potential in the giant squid axon. The intracellular membrane potential typically builds until a threshold is met, which is around –55mV in the giant squid axon. At the threshold voltage, an action potential is generated. For subsequent action potential firings, however, the threshold value changes. Describing these threshold variations requires knowledge of molecular processes controlling the conductance channels that trigger action potentials. This knowledge is not yet understood enough to qualitatively describe the electrical signal behavior [MacG91].

The general shape of the action potential is caused primarily by significant increases in P_{K+} and P_{Na+}, where P_{K+} is a smoother, more gradual increase than P_{Na+}. Both return close to the original values: P_{Na+} within about 1 ms and P_{K+} within about 2 ms (see Figure 2.3.5-1). Both are positive ions, but $[K]_i > [K]_o$ and $[Na]_i < [Na]_o$, so the effect on E are opposing [Dowl92, Kand81]. If P_{Na+} increases, more positive ions will flow from the outside to the inside, raising the potential between the inside and outside, denoted as E_m. However, if P_{K+} increases, more positive ions will flow from the inside to the outside, lowering the potential between the inside and outside.

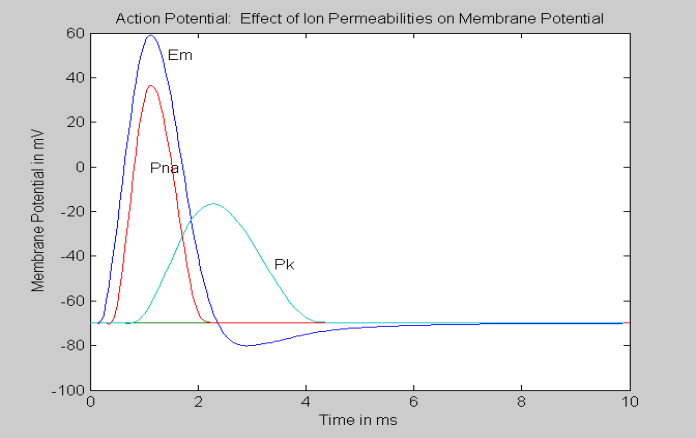


Figure 2.3.5–1. Action Potential and Ion Permeabilities.

After input signals cause E_m to increase past a threshold of about –55mV, a sharp increase in P_{Na} causes E_m to sharply increase to about +60mV. As P_{Na} declines, P_K increases, causing E_m to eventually overshoot the resting potential before settling back there (about –70mV).

The action potential sequence is therefore something like this:

Resting potential:
$E_m \approx$ –70 mV

Cell receives input:
E_m increases to about –55 mV

Action potential initiated:
Sharp increase in P_{Na+}, further increasing E_m to about +50 mV
More gradual increase in P_{K+} as P_{Na+} decreases, reducing E_m
P_{Na+} returns to resting value, while P_{K+} is still high
E_m reduces to about –80 mV (overshoot) as P_{K+} decreases
Ion concentrations and permeabilities return to resting sate, $E_m \approx$ –70

The stronger the input signal to the cell is, the more frequent the action potentials. If the input is present but very weak, then E_m may settle somewhere between –70 mV and –55 mV with no action potentials. An analogy might be a leaky cup being filled from a faucet — when filled, it is emptied, simulating the action potential. If the incoming water flow is not sufficient, an action potential is never generated.

2.3.6 Axonal Signal Transmission

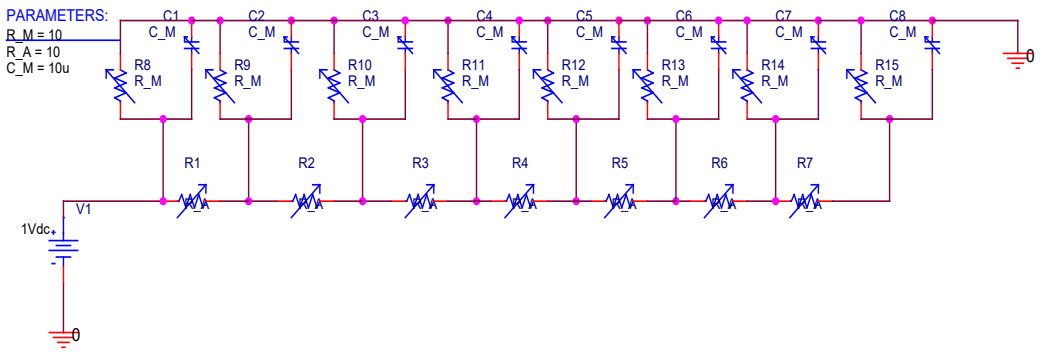
Neurons receive inputs from ionic channels called *dendrites* and transmit (output) signals through their *axons*. Some neurons have no axons and serve to mediate signals by allowing ionic charges to be shared between adjacent neurons. Neurons transmitting action potentials typically have long, conductive axons for the signal transmission. Figure 2.3.6–1 shows a lossy transmission line circuit that simulates the charge-transmission behavior of an axon with membrane capacitance, C_M, membrane resistance, R_M, measured in Ω-cm (longer axons or wider-diameter axons ==> more surface area, less resistance), and axonal resistance R_A, measured in Ω/cm (each unit length can be thought of as a series resistor). The *space constant*, λ, is given as

$$\lambda = \sqrt{\frac{R_M}{R_A}} \quad \Rightarrow \quad V_x = V_0 e^{-x/\lambda}$$

Space Constant Equation

The space constant is analogous to the time constants of RC circuits. For rapidly changing signals, the membrane *time* constant is $R_M C_M$. V_0 here is the DC value at the beginning of the transmission line, and V_x is the value at a distance x from there. Typical values of λ are on the order of 0.1 to 1.0 mm [Kand81]. The space constant is determined from DC values, so C_M is not considered.

Figure 2.3.6–1. Axon Model Circuit



As an action potential propagates down an axon, energy is lost to the series cytoplasm resistance as well as the parallel membrane resistance. The speed of transmission is slowed by the membrane capacitance.

The resistance of a conductor decreases as the cross-sectional area increases. As a result, some species have developed relatively large axons, such as the giant squid axon reaching about 1mm in diameter. The resistance decreases in proportion to the square of the diameter, but the capacitance increases in proportion. The net effect is a decrease in the time constant, $R_M C_M$, resulting in faster transmission. Another biological approach is surrounding the axon with an insulating layer of *myelin*, called the *myelinated sheath*. The result is an increase in the separation of the membrane capacitance charge densities, which reduces the $R_M C_M$ time constant as the capacitance is inversely proportional to the separation distance

Example 2.3.6–1

If the membrane resistance is $R_M = 100\ \Omega$ mm and axonal resistance is $R_A = 10\ \Omega$ /mm, at what length will a DC the signal be down to 10% of its original value?

Solution:

The space constant equation gives

$$\lambda = \sqrt{\frac{R_M}{R_A}} = \sqrt{\frac{100\ \Omega\ mm}{10\ \Omega\ /\ mm}} = 3.16\ m$$

The length where the input, V_0, is down to 10% of its original value is found by

$$0.1V_0 = V_0 e^{-x/\lambda}$$

$$0.1 = e^{-x/3.16}$$

$$\ln 0.1 = -\frac{x}{3.16}$$

$$x = 7.28\ m$$

■

Exercise 2.3.6–1

If the membrane resistance is $R_M = 80\ \Omega$ mm and axonal resistance is $R_A = 15\ \Omega$ /mm, at what length will a DC the signal be down to a) 10% of its original value, b) 5% of its original value, and c) 1% of its original value?

Answers

a) 5.32 mm, b) 6.92 mm, c) 10.64 mm

■

2.3.7 Neuronal Adaptation through Lateral Inhibition

Lateral inhibition is a general phenomenon occurring frequently in layers of interconnected neurons. Cells are electrically coupled so that when one cell fires, it inhibits cells in its neighborhood from firing. The more directly connected a cell is (for example, a nearest neighbor), the greater the inhibition effect. This process is exhibited in the horizontal cell (HC) layer in the retina to provide the spatial-temporal smoothing function [Dowl87]. The HC layer can be modeled as a 2D resistor-capacitor (RC) ladder network, where the RC time constant is the inhibiting force of the network [Koch91].

Lateral inhibition in retina decreases photoreceptor output signal due to activity in nearby photoreceptors; therefore, photoreceptor outputs *adapt to significant changes in the local neighborhood*. In engineering terms, this is considered a *localized automatic gain control* (AGC). A simple AGC applied across the whole image would help prevent image saturation when bright lights are present and help bring out darkened details in the absence of bright lights. In a conventional camera system, a single gain adjustment may be applied to the whole image based on image content. A variation of this is the shutter speed on film that limits the time duration for receiving photons. A high-speed (short integration time) film may be used in the presence of bright lights, while a slow-speed (longer integration time) film may be used in darkened rooms. The difficulty is that such a choice is made across the entire image. It would not be possible to capture both dark and bright contrasts in the same picture.

A *localized* AGC, however, will prevent saturation due to bright intensity sources *and* allow for sufficient detail to bring out dark objects against a dark background. Lateral inhibition in a neuronal layer effectively varies the ACG across the processing plane based on the average activity in the localized area. This is accomplished in the first horizontal cell layer in the retina.

The second laterally-connected retinal cell layer, the amacrine cells, add further lateral inhibition to retinal processing which causes ganglion cell outputs in optic nerve to adapt to motion in the image stream. Motion information is sent initially, but then inhibited by this layer. Similar adaptation to transient signals is observed in the auditory system, as attention is quickly drawn to the onset of a sound, but then suppressed as the neuronal inhibitory signals adapt to this stimulus.

2.3.8 A Circuit Model of a Neuron in Equilibrium

The ionic permeabilities (P_{K+}, P_{Na+}, and P_{Cl-}) regulate how slow or fast ions can move from inside-to-outside or outside-to-inside through a neuron's cell membrane. These parameters can be modeled as conductances, such as G_{K+}, or resistances, such as $R_{K+} = 1/G_{K+}$. The Nernst potential associated with each ion can be modeled as an independent source. Since $[Cl^-]_i \approx [Cl^-]_o$ the Nernst potential for Cl^- is essentially zero. This is due in part by the fact that P_{Cl-} tends to change to keep the concentrations about the same.

Figure 2.3.8–1 shows a simplified model of a neuron in equilibrium. The independent sources represent the Nernst potentials due to the concentration differences between inside and outside the cell. To determine the resting membrane potential, we use Kirchoff's Voltage Law (KVL) to solve for the inside with respect to the outside (ground).

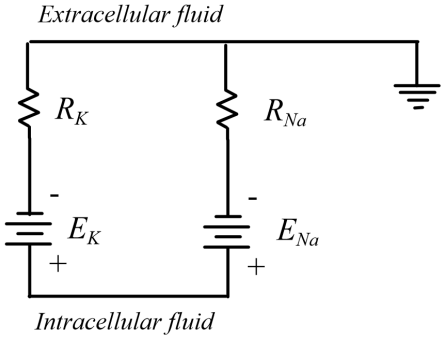


Figure 2.3.8–1. A simplified model of a neuron in equilibrium.

Example 2.3.8–1

Using the model shown in Figure 2.3.8–1 where R_{K+} = 2MΩ, R_{Na+} = 1MΩ, E_{K+} = -75mV, and E_{Na+} = +55mV determine the intracellular fluid potential with respect to the extracellular fluid.

Solution:

Letting I_K be the current upward through R_K and I_{Na} be the current upward through R_{Na},

$-75\text{mV} + I_K R_K - I_{Na} R_{Na} - 55\text{mV} = 0$, and $I_K = -I_{Na}$

$\Rightarrow I_K(R_K + R_{Na}) = 130\text{mV}$

$\Rightarrow I_K = 43.3\text{nA}$

$\Rightarrow V_m = -75\text{mV} + I_K R_K = -75\text{mV} + (43.3\text{nA})(2\text{M}) = \underline{11.67\text{ mV}}$

Check: $-V_m = +55\text{mV} + I_{Na} R_{Na} = +55\text{mV} + (-43.3\text{nA})(1\text{M}) = \underline{11.67\text{ mV}}$

■

Figure 2.3.8–2 shows the same model with dependent current sources representing the Na-K Pump. Sometimes the K^+ ions pumped in are not at the same rate as the Na^+ ions pumped out. However, if the neuron is in a steady-state condition and $[K^+]_i$ and $[Na^+]_i$ are constant, then computations are simplified as

$$I_K = -I_{K\text{-}Pump} \qquad \text{and} \qquad I_{Na} = -I_{Na\text{-}Pump}$$

To solve for the resting potential, V_m, observe these two equations shown above that can be combined with these current relationships:

$$V_m = E_k + I_K R_K \qquad \text{and} \qquad V_m = E_{Na} + I_{Na} R_{Na}$$

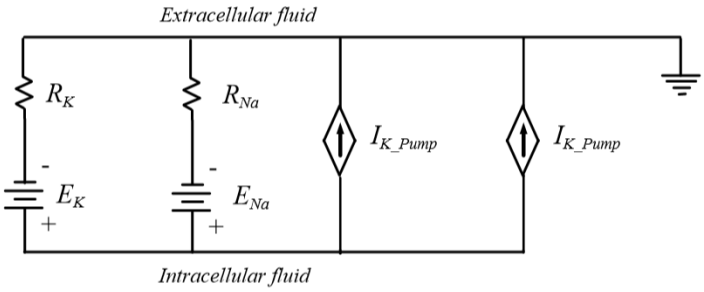


Figure 2.3.8–2. A model of a neuron in equilibrium with ion-pump current sources

This one includes dependent current sources representing organic Na_K Pump.

Example 2.3.8-2

In Figure 2.3.8-2 let I_K be the upward current through R_K and I_{Na} be the upward current through R_{Na}. Calculate V_m, I_K, and I_{Na} if E_k = -75 mV, E_{Na} = +55 mV, $P_K = g_K$ = 1μ mho, $P_{Na} = g_{Na}$ = 0.2μ mho, and 4 Na^+ ions are pumped out for every 3 K^+ ions pumped in. Assume $[K^+]_i$ and $[Na^+]_i$ are constant, and other ionic influences can be neglected.

Solution:

$$R_K = 1/g_K = 1\text{MΩ}$$
$$R_{Na} = 1/g_{Na} = 5\text{MΩ}$$

"...4 Na^+ ions out for every 3 K^+ in..."

$$\Rightarrow \quad 3I_{Na\text{-}Pump} = -4I_{K\text{-}Pump}$$
$$\Rightarrow \quad I_{K\text{-}Pump} = -0.75I_{Na\text{-}Pump}$$

"...$[K^+]_i$ and $[Na^+]_i$ are constant..."

$$\Rightarrow \quad I_{Na} = -I_{Na\text{-}Pump}$$
$$\Rightarrow \quad I_K = -I_{K\text{-}Pump} = 0.75I_{Na\text{-}Pump} = -0.75I_{Na}$$

V_m = -75mV + $I_K R_K$ = +55mV + $I_{Na} R_{Na}$
=> -130mV = $I_{Na} R_{Na}$ -(-0.75I_{Na})R_K
=> I_{Na} = -22.61 nA
=> I_K = (-0.75) I_{Na} = 16.96 nA

V_m = -75mV + $I_K R_K$ = -58.04 mV
Check: V_m = +55mV + $I_{Na} R_{Na}$ = -58.04 mV

■

Exercise 2.3.8-1

Let E_{K+} = -75mV and E_{Na+} = +55mV in the circuit of Figure 2.3.8-1, where I_K is the current upward through R_K and I_{Na} is the current upward through R_{Na}. Calculate V_m for

a) R_{K+} = 200KΩ, R_{Na+} = 5MΩ,
b) R_{K+} = 2.6MΩ, R_{Na+} = 2.6MΩ, and
c) R_{K+} = 5.1MΩ, R_{Na+} = 100KΩ

Based on these values, complete the rest of the following table:

	V_m (mV)
$R_{K+} << R_{Na+}$	_____
R_{K+} = 0.2MΩ, R_{Na+} = 5MΩ	_____
R_{K+} = 2.6MΩ, R_{Na+} = 2.6MΩ	_____
R_{K+} = 5.1MΩ, R_{Na+} = 0.1MΩ	_____
$R_{K+} >> R_{Na+}$	_____

Answers:

	V_m (mV)
$R_{K+} << R_{Na+}$	*~ -75 mV*
R_{K+} = 0.2MΩ, R_{Na+} = 5MΩ	*-70 mV*
R_{K+} = 2.6MΩ, R_{Na+} = 2.6MΩ	*-10 mV*
R_{K+} = 5.1MΩ, R_{Na+} = 0.1MΩ	*52.5 mV*
$R_{K+} >> R_{Na+}$	*~ 55 mV*

■

Exercise 2.3.8-2

In Figure 2.3.8-2 let I_K be the upward current through R_K and I_{Na} be the upward current through R_{Na}. Calculate V_m, I_K and I_{Na} in the model in Figure 2.3.8-2 if E_k = -75 mV, E_{Na} = +55 mV,

$P_K = g_K$ = 5μ mho, $P_{Na} = g_{Na}$ = 0.2μ mho, and 3 Na^+ ions are pumped out for every 2 K^+ ions pumped in. Assume $[K^+]_i$ and $[Na^+]_i$ are constant, and other ionic influences can be neglected.

Answers:

V_m = -71.6mV, I_K = 16.97nA, and I_{Na} = -25.32nA

■

Exercise 2.4-3

Rework Exercise 2.3.8-2 assuming 5 K^+ ions are pumped in for every Na^+ ion pumped out:

In Figure 2.3.8-2 let I_K be the upward current through R_K and I_{Na} be the upward current through R_{Na}. Calculate V_m, I_K and I_{Na} in the model in Figure 2.3.8-2 if E_k = -75 mV, E_{Na} = +55 mV,

$P_K = g_K$ = 5μ mho, $P_{Na} = g_{Na}$ = 0.2μ mho, and 5 K^+ ions are pumped in for every Na^+ ion pumped out. Assume $[K^+]_i$ and $[Na^+]_i$ are constant, and other ionic influences can be neglected.

Answers:

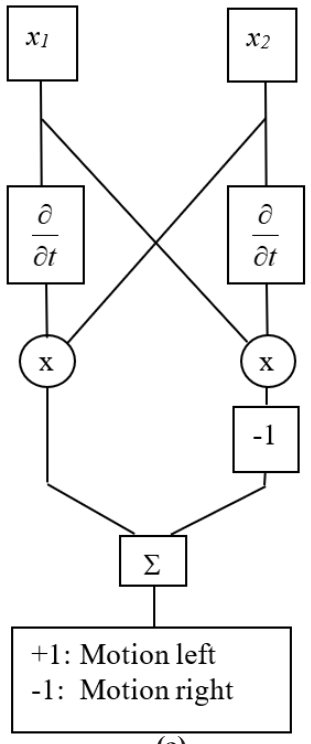
V_m = -53.3 mV, I_K = 108.3 nA, and I_{Na} = -21.67 nA

■

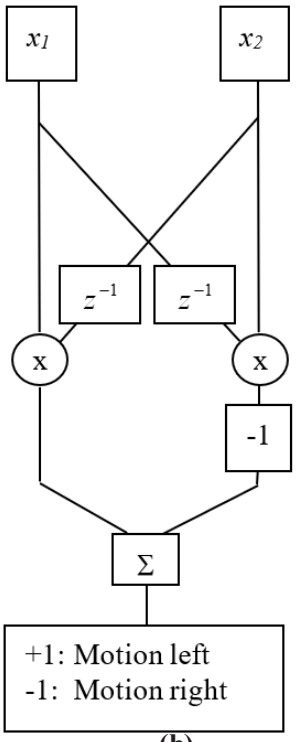
2.3.9 Neuronal Motion Detection

Inputs from adjacent neurons can be connected in a way that provides motion detection. Figure 2.5–1(a) and (b) show two versions of a *Hassenstein-Reichardt* motion detection model [Zorn90, Hass56]. The two inputs, $\underline{x}_1$ and $\underline{x}_2$, represent outputs from two adjacent receptors in a sensory system. In the first instance, a time derivative of one input is multiplied by the value of the adjacent receptor. In the second instance, a delayed version of one input is multiplied by the value of the adjacent receptor. In both instances, the outputs of both products are compared: If equal, they cancel each other in the summation. Otherwise, the result is either positive or negative, depending on the direction of the object.

In Figure 2.3.9–1(c) a three-element binary block is moved right (upper half) and then left (lower half). The input is a ten-element array, and the block is seen by the pattern of ones. The motion detector output is shown, followed by the interpreted results. Both models (a) and (b) result in the same output and interpreted results. The example here is binary, so that the results are very clean. For more realistic values, thresholds would need to be established to reduce effects of noise.



(a)



(b)

x_1	x_2	output	Result
0 1 1 1 0 0 0 0 0 0		0	No Motion
0 0 1 1 1 0 0 0 0 0		0	No Motion
0 0 0 1 1 1 0 0 0 0		-1	Motion Right
0 0 0 0 1 1 1 0 0 0		0	No Motion
0 0 0 0 0 1 1 1 0 0		-1	Motion Right
0 0 0 0 0 0 1 1 1 0		0	No Motion
0 0 0 0 0 0 0 1 1 1		0	No Motion
0 0 0 0 0 0 0 0 1 1		0	No Motion
0 0 0 0 0 0 1 1 1 0		0	No Motion
0 0 0 0 0 1 1 1 0 0		0	No Motion
0 0 0 0 1 1 1 0 0 0		1	Motion Left
0 0 0 1 1 1 0 0 0 0		0	No Motion
0 0 1 1 1 0 0 0 0 0		1	Motion Left
0 1 1 1 0 0 0 0 0 0		0	No Motion
1 1 1 0 0 0 0 0 0 0		0	No Motion
1 1 0 0 0 0 0 0 0 0		0	No Motion

(c)

Figure 2.3.9–1. Hassenstein-Reichardt Motion Detectors.

Two implementations shown: (a) in-channel differentiators, and (b) cross-channel delays. (c) MATLAB simulation results are the same for both implementations.

Exercise 2.3.9-1

Give the expected output (Right, Left or No Motion) of a two-element Hassenstein-Reichardt motion detector given the following two input sequences; assume all previous values are zero. Either model (in-channel derivatives or cross-channel delays) should give the same results:

Sequence 1

x_1	x_2	Output
0	0	No Motion
0	0	_____
1	1	_____
0	0	_____
1	0	_____
0	1	_____
1	0	_____
0	0	_____

Sequence 2

x_1	x_2	Output
0	0	No Motion
1	0	_____
0	1	_____
0	0	_____
0	1	_____
1	1	_____
1	0	_____
0	0	_____

Answers:

Sequence 1

x_1	x_2	Output
0	0	No Motion
0	0	*No Motion*
1	1	*No Motion*
0	0	*No Motion*
1	0	*No Motion*
0	1	*Motion Right*
1	0	*Motion Left*
0	0	*No Motion*

Sequence 2

x_1	x_2	Output
0	0	No Motion
1	0	*No Motion*
0	1	*Motion Right*
0	0	*No Motion*
0	1	*No Motion*
1	1	*Motion Left*
1	0	*Motion Left*
0	0	*No Motion*

Chapter 2 Questions

1. Nothing in the universe is linear, and everything varies with time. Why do we study *linear time-invariant* systems if they do not exist?

2. What are the primary physical sensor types most useful for reverse-engineering? How do the primary senses (touch, taste, smell, vision, and hearing) fit into these types?

3. What are the three basic types of stimulus reception in biological sensory systems?

4. What are the three basic maps of sensory receptive fields found in the brain?

5. What are the similarities between natural and artificial neural networks?

6. If neurons are so much slower than transistors, how could there be promise of significant performance improvement for computers built with diodes and transistors that model neuronal behavior?

7. If biological systems are constantly maturing and adapting, why is it beneficial to study the structure and function of the neuronal system of a mature adult animal (or human).

8. What is an immediate environmental adaptation in the human vision system?

9. Compare and contrast charge and steady-state charge neutrality in neurons and transistors.

10. How do the membrane resistance and axonal resistance affect transmission ability of an action potential down an axon? What other factors can improve transmission? That is, what other factors increase the length of the spatial constant?

11. If a layer of cells exhibits lateral inhibition and a single neuron fires (produces an action potential), what happens to adjacent cells that are connected to this cell? What happens to cells that are connected but are farther away?

■ 12. How are signals typically processed in neuronal layers? Examples may include the neuronal layers of the retina or the brain.

13. How is the strength of a signal measured when encoded as action potentials of the same peak value (around +55 mV)?

14. Action potentials are triggered when the intracellular fluid potential exceeds a threshold. How is it that for a steady input the output firing rate (frequency of action potentials) adapts from an initial firing rate to a slower firing rate?

15. What controls the rate of adaptation?

Chapter 2 References

[Dowl87] Dowling, J. E., *The Retina: An Approachable Part of the Brain*, Harvard University Press, Cambridge, Massachusetts, 1987.

[Dowl92] Dowling, J. E., *Neurons and Networks–An Introduction to Neuroscience*, Harvard University Press, Cambridge, Massachusetts, 1992.

[Hass56] Hassenstein, B. and Reichardt, W., System-theoretische Analyse der Zeit-, Reihenfolgen und Vorzeichenauswertung bei der Bewegungsperzeption des Russelkafers Chlorophanus. *Zeitsschrift fur Naturforschung, Teil* B, Vol. 11, pp. 513–524, 1956.

[Hech90] Hecht-Nielsen, R., *Neurocomputing*, Addison-Wesley, New York, 1990.

[Horen96] Horenstein, M., *Microelectronic Circuits and Devices,* Pearson, 1996.

[Kand81] Kandel, E. R. and Schwartz, J. H., *Principles of Neural Sciences*, Elsevier/North-Holland, New York, 1981.

[Koch91] Koch, C., "Implementing early vision algorithms in analog hardware: An introduction", in Mathur, B. P., and Koch, C., Eds., *Visual Information Processing: From Neurons to Chips*, *Proc. of the SPIE,* Vol. 1473, 1991.

[MacG91] MacGregor, R. J., *Neural and Brain Modeling*, Academic Press, Inc., New York, 1987.

[MatLab] MatLab, which stands for "MATrix LABoratory", is a trademark of the computational software product developed by the Mathworks, Inc.

[Vass95] Vassileu, A., Mitou, D., and Manahilou, V., "Grating detection and identification dissociated by pattern adaptation", *Spatial Vision*, Vol. 9, No. 2, pp. 221–234, 1995.

[Warner77] Warner, G. F., *The Biology of Crabs*, Van Nostrand Reinhold Company, NewYork, ISBN: 0-442-29205-8, 1977.

[Zorn90] Zornetzer, S., Davis, J., and Lau, C., Eds., *An Introduction to Neural and Electronic Networks*, Academic Press, Inc., ISBN: 0-12-781881-2, 1990.

Chapter 3: PHOTO-SENSORY SYSTEMS

Biological sensory systems perform energy-efficient and computationally elegant algorithms to accomplish tasks like those required of certain engineering applications. Animals and some engineered systems have the capacity for limited movement within the natural environment in response to sensory stimuli. For example, consider a front-end seeker on a missile designed to autonomously seek and hit a specified target. The missile needs to be guided to a target seen by a seeker with background sensory noise; this requirement is like that of a dragonfly searching and acquiring smaller flying insects. Tasks common to both systems include navigating and guiding the system within the natural environment, detecting, identifying, and tracking objects identified as targets, efficiently guiding the system to the targets, and then intercepting these targets.

This part is about photo-sensory systems, or vision, which involves the conversion of photonic energy into electronic signals. These signals are subsequently processed to extract pertinent information. The primary emphasis will be on vision computational models based on the primate vision system since much study has been made in this area. We begin with some vision principles common across many species within the animal kingdom. Then the structure and function of natural vision systems is investigated, with emphasis on information processing first within invertebrates (specifically arthropods) and then within vertebrates (specifically primates). Engineering application examples that leverage natural vision concepts follow.

3.1 Natural Photo-sensory Systems

Passive means the sensor observes natural stimuli that might be available within the environment, while *active* implies the sensor sends stimuli out and observes the response from the environment. Physical sensors in the animal kingdom include *photo-sensory*, such as passive vision systems processing photons, *mechano-sensory*, such as passive sonar (audition), active sonar (bats, dolphins, whales), passive compression (touch) and active compression (insect antennas), and *chemo-sensory*, such as gustation (taste) and olfaction (smell). This chapter will focus on passive photo-sensory vison systems.

3.1.1 Common principles among natural photo-sensory systems

A *photon* is the wave-particle unit of light with energy $E = h\upsilon$, where h is Plank's constant and υ is the electromagnetic frequency. The energy per time (or space) is modeled as a *wavelet* since it satisfies the general definition of having a beginning and ending and unique frequency content. Information contained in the frequency and flux of photons is photonic information, which gets converted into electronic information coded in the graded (or analog) neural ionic voltage potentials or in the frequencies of action potentials

Biological systems can be divided into *vertebrates*, such as mammals and reptiles, and *invertebrates*, such as insects. Animals collect and process information from the environment for the determination of subsequent action. The many varied species and associated sensory systems in existence reflect the wide range of environmental information available as well as the wide range of biological task objectives.

Commonality of Photo-reception and Chemo-reception

Photo-reception is made possible by the organic chemistry of *photopigments*, which initiate the visual process by capturing photons of light. Photopigments are composed of a form of Vitamin A called *retinal* and a large protein molecule called *opsin*. Opsins belong to a large family of proteins which include olfactory (sense of smell) receptor proteins. Odorant and tastant molecules attached themselves to a special membrane receptor, causing a sequence of molecular reactions eventually resulting in neuronal signaling. Photopigment molecules are like these chemo-sensory membrane receptors with retinal serving as the odorant or tastant already attached. The incoming photon of light gives the molecule enough energy to initiate a chain reaction like that in chemo-sensory reception when an odorant or tastant molecule come in contact with the receptor. As a result, the photo-reception process is really a simplified form of the chemo-reception process. A photo-sensory (or visual) system begins by converting the photonic stimulus into a chemical stimulus (photopigments) and the remaining information processing of the visual system is that of a chemo-sensory system.

Curvature and Reflection

The two primary eye designs are the *vesicular* (containing a cavity) eye found in vertebrates and certain mollusks and the *compound* eye found in arthropods. Figure 3.1–1 shows the concave nature of the vesicular eye and the convex nature of the compound eye. Images in biological systems are formed on a curved sheet of photoreceptors, called the *retina*. In a similar way, cameras form images on a sheet of photographic film, where the film is flat instead of curved. The ancient sea-going mollusk *Nautilus* has the concave retina structure with a pinhole aperture, which creates an inverted image with no magnification. Most concave retinas (vertebrates, etc.) depend on the refraction of light through a larger aperture. The *lens* serves this purpose. A larger aperture is needed to allow more photonic flux to enter the reception area to ensure sufficient energy is available to stimulate photoreceptors, and refraction through the lens and eyeball fluid (*vitreous humor*) serves to compensate for the otherwise blurred view of the environment as the aperture is increased.

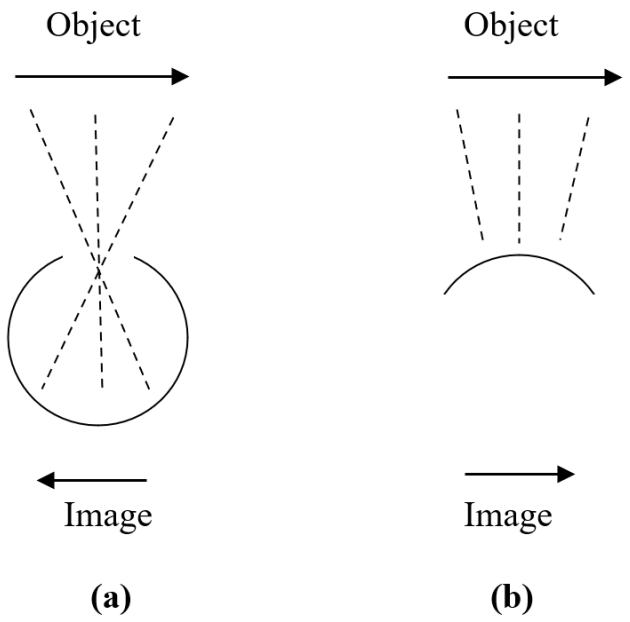


Figure 3.1–1

(a) Concave retina of the vesicular eye of many vertebrates and mollusks.
(b) Convex retina of arthropod compound eye.

Physical properties of reflection are also used in the eye designs of scallops and certain fish and mammals. Some of the purposes of designs based on reflection are not known (scallops), but other vision system designs exploit reflection in nocturnal (night-time low-light level) conditions. For example, night-hunting by certain mammals is augmented by the fact that a photon of light has twice as much a chance of being captured by the same photoreceptor as the light passes through a second time after being reflected. A special reflective tissue (*tapetum lucidum*) behind the retina gives this advantage in nocturnal conditions. This reflection can be observed when shining a light (flashlight, headlight) toward the animal and it is looking back.

The photoreceptors are typically long and cylindrical cells containing photopigments arranged in many flat disc-shaped layers. This design gives a small angular reception area, leading to sufficient spatial *acuity*, while providing many opportunities for the incoming photon to be captured by the photopigment.

Optical Imperfections

There are several imperfections that are dealt with in natural vision systems. Some of these include spherical aberration, chromatic aberration, and diffraction. Natural vision system parameters typically represent an optimal balance of the effects of these imperfections. *Spherical aberration* is caused by light coming into focus at a shorter distance when coming through the periphery of the lens than from the center. *Chromatic aberration* is caused by the dependency on wavelength of the index of refraction: The shorter the wavelength, the greater the amount of refraction. This means that if the blue part of the image is in focus, then the red part of the image is slightly out of focus. The optical properties of the available biological material do not allow for perfect compensation of these effects. For example, to correct for spherical aberration requires a constant decrease in the cornea index of refraction with distance from the center. Since the molecular structure of the cornea is constant, this is not possible. The general shape, however, of the primate eye is slightly aspherical, which minimizes the effects of spherical aberration. As the primate eye changes shape with age, these aberrations are corrected by external lenses (eyeglasses).

The third imperfection is caused by *diffraction*. Diffraction is a geometrical optics phenomenon resulting from the edge effects of the aperture. When combined with spherical and chromatic aberration, the result is a spatial frequency limit on the image that can be mapped onto the retina. This limit is typified by the angular distance that two separate point sources can be resolved, called *angular acuity. Spatial acuity* refers to the highest spatial frequency that can be processed by the vision system. The displacement between photoreceptors in highly evolved species is typically the distance represented by the angular acuity. Any further reduction in distance is not practical as there would be no advantage concerning image information content.

Another consideration is *contrast sensitivity*, which is how sensitive two separate photoreceptors are to varying levels of photon flux intensity. In biological systems, the information forwarded is frequently a difference in contrast between two adjacent photoreceptors. If the photoreceptors are very close, then the difference will never be great enough to show a relative contrast since edges in the image are already blurred due to the aforementioned imperfections. The photoreceptor spacing in the retina is on the order of the Nyquist spatial sampling interval for frequencies limited by these imperfections. In the adult human retina, this turns out to be about 120 million photoreceptors: about 100 million *rods*, which are very sensitive and used in nocturnal conditions, and about 20 million *cones*, which come in three types and provide color information in daylight conditions.

Visual Information Pathways

Receptive fields for the various sensory systems are mapped to specific surface regions of neuronal tissue (such as retina, brain, and other neuronal surfaces). Due to the connectivity, several *pathways* are usually observed. For example, one photoreceptor may be represented in several neurons that are transmitting photonic information to the brain. One neuron may represent the contrast between that particular photoreceptor and the most adjacent ones. This would be an example of a *parvocellular pathway* neuron (*parvo* means small). Another neuron may represent the contrast between an average of that photoreceptor and the most adjacent ones, and an average of a larger region centering on that photoreceptor. This would be an example of a *magnocellular pathway* neuron (*magno* means large). As it turns out, the names come from the relative physical size of these neurons, and they happen to also correspond to the size of the receptive field they represent. Parvocellular and Magnocellular pathways are common among many species, for example, both humans (and other mammals) and certain arthropods.

Connectivity and Acuity

There is a balance between *temporal acuity*, which is the ability to detect slight changes in photonic flux in time, and *spatial acuity*, which is the ability to detect slight changes between two adjacent objects whose images are spatially separated on the retina. As receptors are more highly interconnected, there is better temporal acuity due to the better photon-integrating ability of the aggregate. Receptors that are not highly interconnected exhibit better spatial acuity.

To illustrate this concept, consider a steady photonic flux represented by 1 photon per 10 photoreceptors per unit of time. On average, each photoreceptor would receive 1 photon every 10 units of time. If this incoming photon rate changed to 2 photons per 10 photoreceptors, then the output of a single photoreceptor would have to be monitored for a duration of 10's of units of time to detect an average increase in photon flux. If an aggregate of 100 photoreceptor cells were integrated, and if the photonic flux were uniformly distributed, then the total output would jump from 10 photons to 20 photons, which might be noticeable at the very next unit of time. The result is that the animal will be able to detect slight changes in photonic flux much better if the cells are highly connected, while the ability to distinguish between two adjacent small objects would deteriorate. Thus, *a higher connectivity results in sharp temporal acuity at the cost of spatial acuity*.

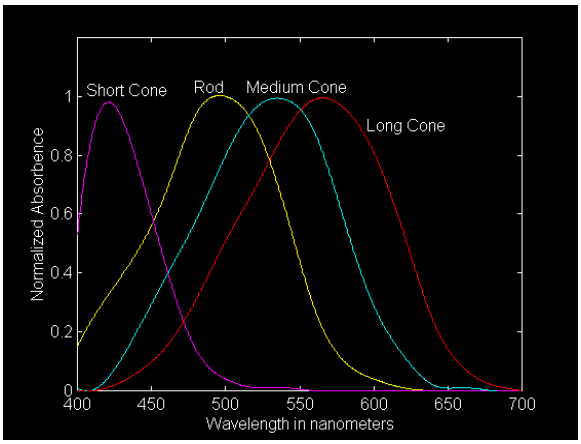
Coarse Coding

Coarse coding is the transformation of raw data using a small number of broadly overlapping filters. These filters may exist in time, space, color, or other information domains. Biological sensory systems tend to use coarse coding to accomplish a high degree of acuity in sensory information domains. For example, in each of the visual system *information domains* (space, time, and color, or chromatic) we find filters that are typically few in number and relatively coarse (broad) in area covered (or bandwidth): There are essentially only four chromatic detector types, whose spectral absorption responses are shown in Figure 3.1–2, three temporal channels, and three spatial channels. Neurons in the retina receiving information from the photoreceptors are connected in such a way that we can observe these spatial, temporal, and chromatic information channels in the optic nerve.

Coarse coding can take on many different forms, and one coarsely coded feature space may be transformed into another. For example, within the color channels of the vision system we find a transformation from broad-band in each of the three colors at the sensory level to broad-band in color-opponent channels at the intermediate level. Other interesting examples of coarse coding include wind velocities and direction calculation by cricket tail sensors and object velocity calculations with bursting and resting discharge modes of neuronal aggregates in the cat superior colliculus.

The responses of vision system rods and cones must be broad in scope to cover their portion of the data space. For example, in daytime conditions only the three cone types have varying responses. As a minimum each type must provide some response over one-third of the visible spectrum. Each detector type responds to much more than one-third of the visible spectrum. Since a single response from a given detector can result from one of many com-

binations of color and intensity, the value by itself gives ambiguous local color and intensity information. If the response curve was very narrow band, then any response is the result of a particular frequency, and the value of the response would reflect its intensity. However, many of these detectors would be required to achieve the wide range (millions) of colors we can perceive. It is not practical to have each of many narrow-band detectors at each spatial location. The natural design is optimized to allow for many colors to be detected at each location while minimizing the neuronal hardware (or "wet-ware") requirements.



Ultraviolet *Visible Spectrum* *Infrared*

Figure 3.1–2. Photon Absorption Curves.

These curves are peak-normalized response curves of the three known photopigments in primate retina. The responses overlap considerably. This figure was produced using Matlab curve-fitting to the selected points in measured biological tissue.

3.1.2 Arthropod vision system concepts

Although there are millions of species within the animal kingdom, there are relatively few photo-receptor design concepts that have stood the test of time, such as the arthropod compound eye. There are some interesting similarities between the vision systems of the insect phyla and primates. For example, both map incoming light onto an array of photoreceptors located in a retina. Both exhibit distinct post-retina neuronal pathways for what appears to be spatial and temporal processing.

Of course, there are some key differences between insect and primate vision systems. Insects have non-movable fixed-focused optics. They are not able to infer distances by using focus or altering gaze for object convergence. The eyes are much closer together, so that parallax cannot be used to infer distances either. The size is much smaller, and the coverage is in almost every direction so that the overall spatial acuity is much worse than primates. As a result, navigation appears to be done more by relative image motion than by any form of object detection and recognition [Srini02].

Arthropod Compound Eye

The arthropod *compound eye* is a convex structure. The compound eye is a collection of individual *ommatidia*, which are complex light-detecting structure typically made up of a corneal lens, crystalline cone, and a group of photosensitive cells. Each ommatidium forms one piece of the input image so that the full image is formed by the integration of all ommatidia. There are three basic designs for integrating ommatidia into a composite image:

1) *Apposition*. Each ommatidia maps its signal onto a single photoreceptor.

2) *Superposition*: Several ommatidia contribute to the input signal for each photoreceptor

3) *Neural superposition*: Not only are the photoreceptor inputs a superposition of several ommatidia, but neurons further in the processing chain also receive their inputs from several photoreceptor outputs.

Apposition eyes form relatively precise images of the environment. This design is common among diurnal (daytime) insects. Superposition eyes are common among nocturnal (night-time) and crepuscular (twilight) insects. In conditions of low light levels, the superposition design allow for greater sensitivity since light from several ommatidia are focused onto a single photoreceptor. The greater sensitivity of the superposition eye comes at a cost of spatial acuity since image detail is shared by neighboring pixels. This is an example of "higher connectivity results in sharp temporal acuity at the cost of spatial acuity" explained earlier. The neural superposition eye is found in the dipteran (two-winged) fly. This design allows for further processing to compensate for the loss of spatial acuity, resulting in both good spatial acuity and sensitivity.

The superposition eye has greater sensitivity to changes in photonic flux because of the higher degree of connectivity of the ommatidia to a single photoreceptor. In a similar way, the primate rod system is highly interconnected, which results in a high degree of temporal sensitivity. The primate photoreceptors are divided into *rods* and *cones*, named for the shape of the outer photopigment-containing segment. Certain cone cells are also highly interconnected, bringing better sensitivity to temporal changes.

Scanning Eyes

A few mollusks and arthropods have developed a scanning mechanism for creating a visual image of the external environment. A narrow strip of photoreceptors is moved back and forth to generate the complete image. Certain sea snails have retinas that are 3 to 6 photoreceptors wide and 400 photoreceptors long. The eye scans 90°, taking about a second to scan up, and about a fourth of a second to return down [Smith00].

Mantis shrimp contain 6 rows of enlarged ommatidia in the central region of the compound eye. The larger ommatidia contain color visual pigments that can be used to further investigate an object of interest by scanning with these central photoreceptors. This allows the shrimp to use any color information in the decision process[Smith00].

Certain jumping spiders contain retinas 5 to 7 photoreceptors wide and 50 photoreceptors long. The spider normally scans from side but can rotate the eye to further investigate a particular object of interest. The lateral (additional) eyes on this spider contain highly interconnected photoreceptors for detecting slight rapid movements. Once detected, the attention of the primary eye can be directed to the newly detected object. This process is analogous to primate vision, where the more periphery cells are highly connected and the central area (the *fovea* to be discussed later) are more densely packed and not so interconnected. A sharp movement in the periphery causes a primate to rotate the eyes to fixate on the source of the movement. Once fixated, the higher spatial acuity of the central area can be used to discern the spatial detain of the new object of interest [Smith00].

3.1.3 Primate vision systems

Early vision can be defined as the processes that recover the properties of object surfaces from 2D intensity arrays. Complete vision would be the process of using early vision information to make some decision. The focus in this section is on vertebrate vision information pathways that begin in the retina and terminate in cortical processing stages. *Cortical* comes from *cortex*, which is used to describe the part of the brain where sensory system information is processed. Vision is processed in the *primary visual cortex,* hearing is processed in the *auditory cortex,* and touch is processed in the *somatosensory cortex.* Many of these concepts are also common in insect vision.

Figure 3.1–3 shows the relevant parts of the primate eye. Photonic energy is first refracted by the cornea and further by the lens and the *vitreous humor*, which fills the optics chamber. The retina covers most of the inner portion of the eye and serves as the first vision processing stage. Approximately 120 million photoreceptors are encoded into about 1 million axons that make up the optic nerve.

Figure 3.1–4 shows the other basic components of the primate vision system. A projection of the 3D environment is mapped onto the 2D sheet of neuronal tissue called the *retina*. The primate retina is composed of several layers of neurons, including *photoreceptor, horizontal, bipolar, amacrine,* and *ganglion* cell layers to be discussed in more detail later. The information is *graded*, which basically means analog to electrical engineers, until it reaches the axon (output) of the ganglion cell layer. The graded potential signaling is replaced by action potential signaling through the optic nerve. Upon reaching the *optic chiasm*, the right side of both retinas (representing the left side of the visual field) are mapped to the right side of the brain, and the left side of both retinas (right side of visual field) to the left side of the brain.

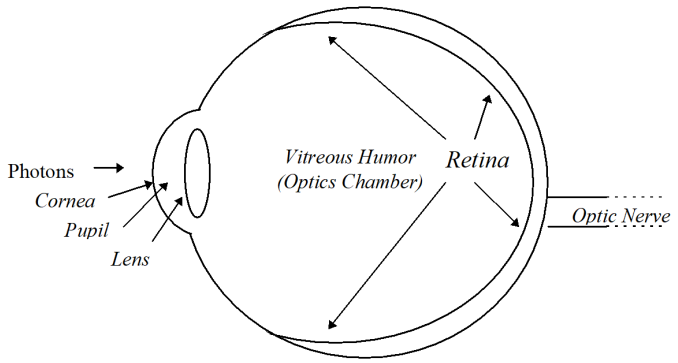


Figure 3.1–3. Simplified anatomy of the primate eye.

The retina, lateral geniculate nucleus (LGN) and the brain are all composed of layers of neurons. Figure 3.1–4 highlights the LGN whose outer 4 layers are the termination of Parvocellular Pathway (PP) optic neurons and inner 2 layers the termination of Magnocellular Pathway (MP) optic neurons. Both PP and MP signals are *opponent* signals, meaning the signal levels correspond to the contrast between a central receptive field (RF) and a larger surrounding RF which would include responses from neurons not represented by the central RF. *Parvo* (small) and *magno* (large) were names given by anatomists who based the names on the size of the cell bodies. Conveniently, it was later learned that the PP corresponds to smaller RFs (central RF could be one cell) and MP to larger RFs (central RF would be a larger aggregate of cells). In both cases the surrounding RF would be larger than the central RF. There is duality in the center-surround contrast signals in that some represent the central signal minus the surround ("ON" signals) while others represent the surround signal minus the central ("OFF" signals).

The PP contains color information as the cone response of a single central signal will have a different spectral response from the average response of the surrounding neurons. Some earlier researchers would use *r, g, b* for designating the three cone receptors. But since the spectral absorption curves broadly overlap much of the visible spectrum (as show in Figure 3.1–2) a better notation is *l, m, s* for *long-, medium-,* and *short-wavelength* cone types [DeV88]. We adopt that convention in this book.

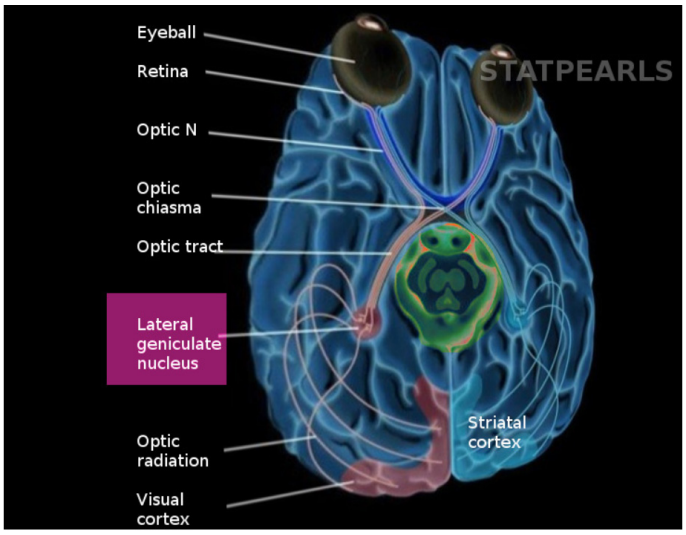


Figure 3.1–4. Basic Components of Primate Vision.

Downloaded August 2020 from
https://www.ncbi.nlm.nih.gov/books/NBK541137/

Spatio-temporal Processing Planes

The retina can be considered a "part" of the brain, as suggested by the subtitle of John Dowling's book The Retina: An Approachable Part of the Brain [Dowl87]. The retina is a multi-layered region of neuronal tissue lining the interior surface of the eye, as shown in Figure 3.1–3. In the early stages of primate central nervous system (CNS) embryonic development, a single neural tube develops two *optic vesicles* with *optic cups* that eventually develop into the retinas for each eye. The physiology (or functioning) of layers of neurons are similar, whether located peripherally in the retina (about 5 layers), in the LGN (about 6 layers), or in the visual cortex (about 10–12 layers). If we can better understand the spatial-temporal-chromatic signal processing that exists in the retinal it will better our understanding of what is also going on in the LGN and the higher processing centers of the visual cortex.

The vision processing mechanics can be best visualized as a series of parallel-processing planes, each representing one of the neuronal layers in the retinal or in the brain, as shown in Figure 3.1–5. Parallel incoming photons are received by the outer segments of the photoreceptors resulting in signals that propagate to the visual cortex in the brain. Each plane of neuronal processing acts upon the image in a serial fashion. However, the processing mechanism cannot be simply described as simple image filters acting on each separate plane. As the energy is propagated through the neuronal layers, the ionic charge spreads laterally across each processing plane. As a result, the output of each processing plane is a combination of the current and historic inputs of the cells in the path as well as the historical input of the adjacent cells.

To adequately model spatial and temporal effects of the neuronal interconnections, each cell in each neuronal processing plane must consider mediation effects of neighboring cells as well as temporal effects of signal degradation in time. One way to model both effects is to apply a 2D spatial filter to each image plane and follow the filter with a leaky integrator, that allows for temporal ionic equilibrium effects.

Information Encoding

Natural vision systems extract space (spatial), time (temporal) and color (chromatic) information to make some decision. Information is often encoded for transmission, for example, from the retina to the LGN. Figure 3.1–6a shows the basic information blocks in the vision system. Figure 3.1–6b illustrates the overall numerical processing elements in each of the various vision processing stages. There is an approximately 100:1 compression of the retina photoreceptors to the optic nerve signals, but an expansion of 1:1000 optic nerve signals to visual cortex neurons. This expansion is known as *optic radiation*. Combining the compression and expansion there is an overall expansion of about 1:10 retinal photoreceptors to visual cortex neurons. As typical in biology, the compression and expansion is quite non-uniform, as there are about 2 optic nerve neurons per photoreceptor in the retina's fovea (very central part of vision), but only 1 optic nerve neuron for about every 400 photoreceptors in the peripheral part of the retina. This unbalance is a consequence of the importance of information in the center-of-gaze.

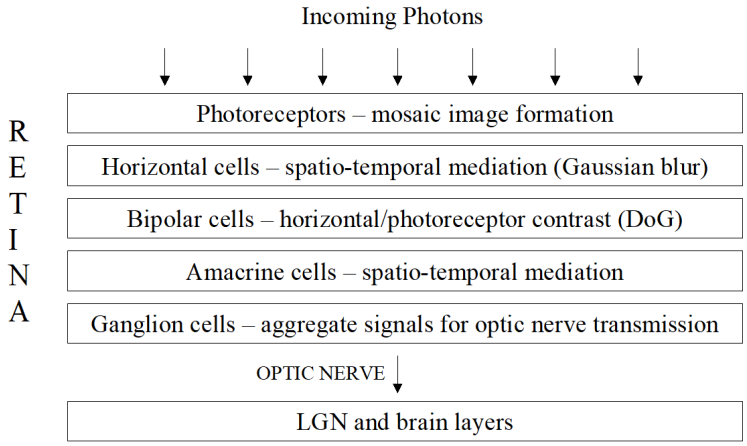


Figure 3.1–5. Planar/serial duality of vision processing.

Information in the form of cellular ionic potentials is filtered spatially as it propagates through the various layers of the vision system. Show in detail are the layers of the retina; layers in the LGN and visual cortex continue the spatio-temporal processing.

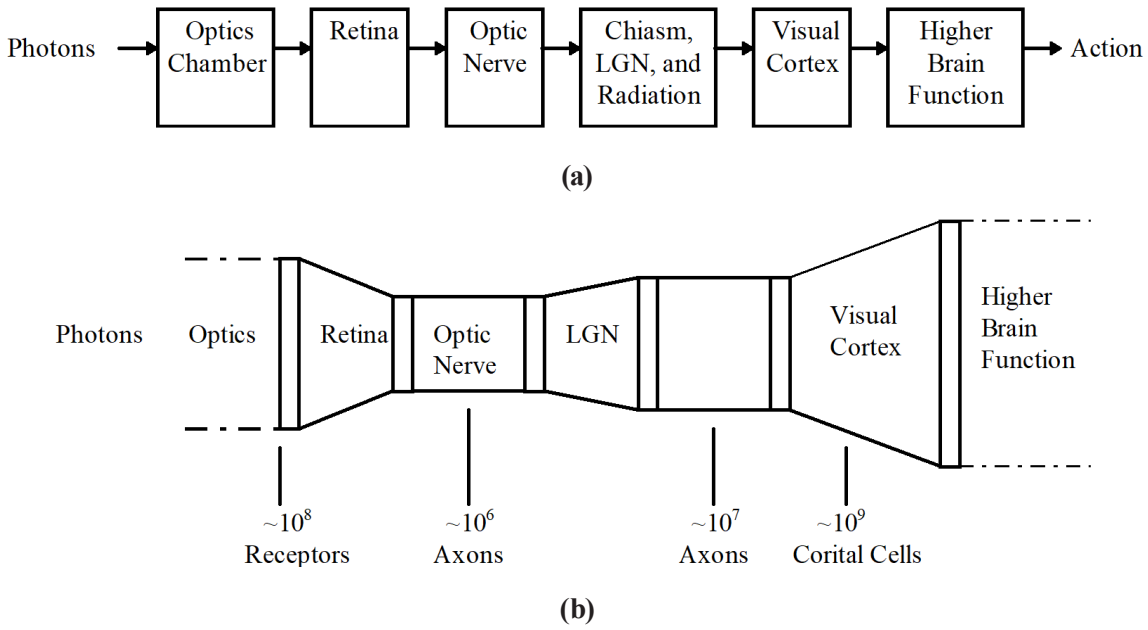


Figure 3.1–6. Functional Vision Blocks (a) and Signal Encoding (b).

Natural vision filtering begins with photonic refraction through the cornea and lens (Figure 3.1–3). Figure 3.1–7 depicts the various cell layers within the retina and a gross approximation of the mathematical function performed by each layer on the incoming imagery. The incoming light then passes through the vitreous humor and retinal cell tissue and is focused onto a *photoreceptor mosaic* surface. The flux within a photoreceptor's receptive region of the retina is averaged to a single output at the *triad synapse* (at the root of the photoreceptor). As a result, the information can be visualized as a mosaic, where each piece represents a single photoreceptor's output.

Photonic energy is converted to electronic charge in the photopigment discs of the photoreceptors (rods and cones). It is believed that the rate of information transfer is proportional to the logarithm of the incoming intensity. The photoreceptors, with the help of a layer of *horizontal* cells, spread the charge in space and time within a local neighborhood of other receptors. Such charge-spreading can be modeled by spatio-temporal gaussian filters. Two separate variances (horizontal and vertical) are required for the spatial 2D filter and another for how the signal degrades in time.

The spread charge and original photoreceptor charge, both of which can be modeled as a gaussian-filtered version of the incoming imagery, are both available at the root of the photoreceptor, at the *triad synapse*. The *bipolar* cells connect to triad synapses and presumably activate signals proportional to the difference between the photoreceptor input and the horizontal cell input. Therefore, the bipolar cell output represents the *difference-of-gaussian* version of the original image.

Spatial edges are detected by two types of bipolar cells, *on*-bipolars and *off*-bipolars, which respond to light and darkness, respectively. The on-bipolar responds if the central receptive field exceeds the surrounding receptive field, while the off-bipolar cells respond if the surrounding receptive field exceed the central receptive field. Temporal edges (rapid changes in photonic flux levels) are detected by *on-off* and *off-on* bipolar cells, which respond to quick decrements or increments in photonic flux, respectively. Corresponding ganglion cells (*on, off*, *on-off*, and *off-on*) propagate amacrine-cell-mediated responses to these bipolar cells.

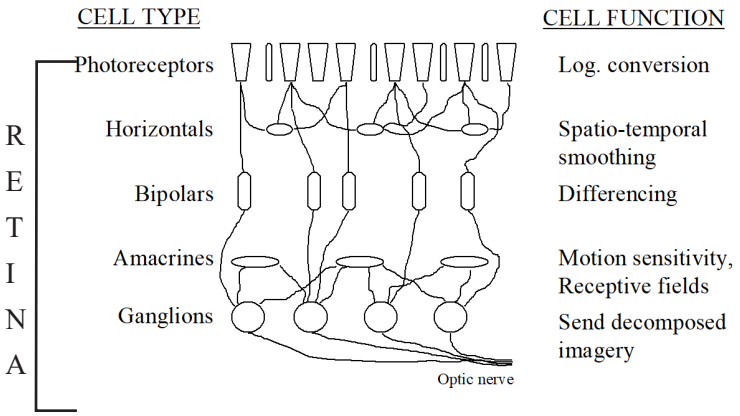


Figure 3.1–7. A generic model of the retina cell layers and corresponding functionality.

The difference signal propagated by the bipolar cells is a consequence of the *lateral inhibition* caused by the connectivity of photoreceptors and horizontal cells. The horizontal cells connect horizontally to numerous photoreceptors at the triad synapse. Horizontal cells only have dendrites, which for other neurons would typically serve as input channels. The dendrites (inputs) for these cells pass ions in both directions, depending how the ionic charge is distributed. The net effect is that adjacent photoreceptors have their information partially shared by this mediation activity of the horizontal cells.

Gap junctions between adjacent photoreceptors influence the photoreceptor charge. The response from a photoreceptor aggregate can be modeled as a spatial-temporal Gaussian with a small variance. The input from the neighboring aggregate of horizontal cells can be modeled with a similar Gaussian with a larger variance. The differencing function results in the *difference-of-Gaussian* (DOG) filter operation, resulting in a center-surround antagonistic receptive field profile. DOG functions and functions of the second derivative of Gaussian, called the *Laplacian-of-Gaussian* (LOG), have been used to model the bipolar cell output.

The analog charge information in the retina is funneled into information pathways as it is channeled from the mosaic plane to the optic nerve. These information channels originate in the retina and are maintained through the optic nerve and to portions of the brain. These include the *rod* channel, initiated by rod bipolars, the *parvocellular pathway* (PP) and the *magnocellular pathway* (MP), the latter two initiated by cone bipolars. Both the PP and the MP exhibit center-surround antagonistic receptive fields. PP cones are tightly connected, responding to small receptive fields, while the MP cones are more loosely connected (together with rod inputs), responding to large receptive fields.

The MP and PP perform separate spatial band-pass filtering, provide color and intensity information, and provide temporal response channels, as illustrated in Figure 3.1-8. A relatively high degree of acuity is achieved in each domain (space, time, and color, or chromatic) from these few filters. The MP is sensitive to low spatial frequencies and broad color intensities, which provide basic information of the objects in the image. The PP is known to be sensitive to higher spatial frequencies and chromatic differences, which add detail and resolution. In the color domain, the PP provides color opponency and thus spectral specificity, and the MP provides color non-opponency and thus overall intensity. In the time domain, the PP provides slowly varying dynamics, while the MP provides transient responses to image dynamics.

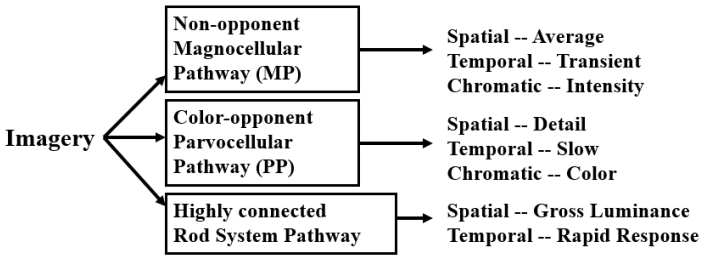


Figure 3.1–8. Natural Vision Information Channels.

The color opponent PP responds to spatial detail, slowly varying image dynamics, and chromatic detail. The color non-opponent MP responds to spatial averages, rapid transients, and intensity variations. The rod system pathway aids overall luminance and provides rapid temporal responses in periphery.

Graded Potential Processing

Retinal information is primarily in the form of graded potentials as it moves from the photoreceptor cell (PC) layer through the retina to the amacrine cell (AC) and ganglion cell (GC) layers. The GC output axons make up the optic nerve, transporting spikes to the LGN. The ganglion axonal signals begin the optic nerve transmission of color, time, and space information to the remaining neuronal organs in the vision pathway. It is typical that localized processing is graded, like an analog voltage level in an RLC circuit, but is pulsed via action potentials when travelling distances, such as from the retina to the LGN, and from there to the superior colliculus and to the visual cortex.

Figure 3.1–9 shows the signal and image processing functions at the various stages of the retina. Figure 3.1–10 shows greater detail of the lower left region of Figure 3.1–9. The spatio-temporal filtering characteristic is due to the connectivity of the first three layers of neurons: photoreceptors, horizontal cells, and bipolar cells.

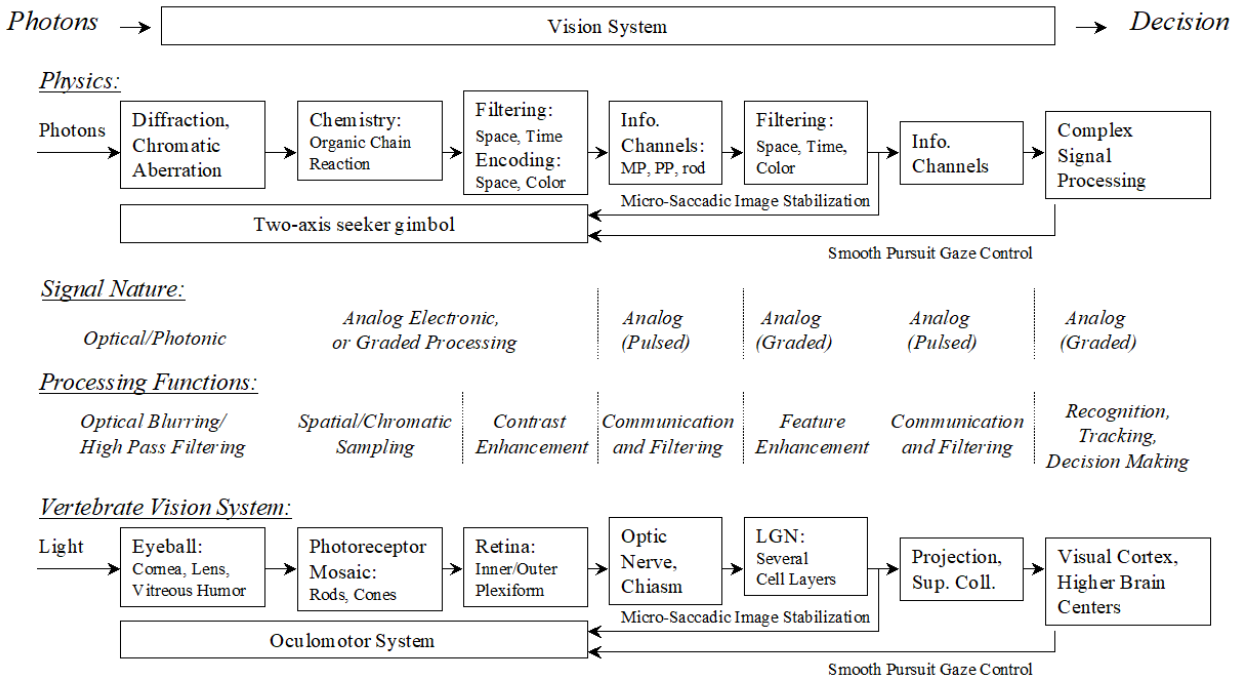


Figure 3.1-9. Physics and Signal Processing in the Vertebrate Vision System.

The overall role (top line) is to process photonic information and make some decision as a result. This figure illustrates vertically the physics, signal nature, and processing functionality with respect to the various organic stages as information moves horizontally (to the right).

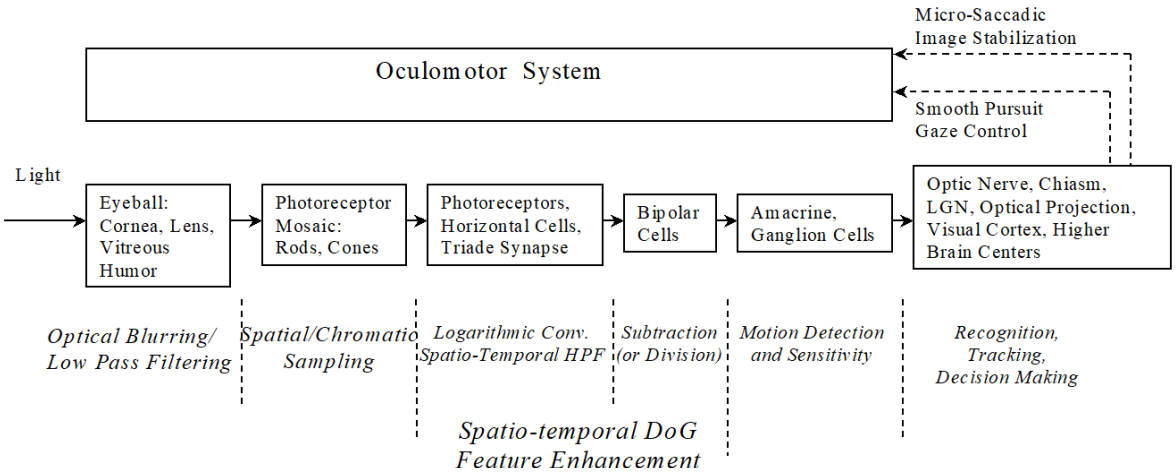


Figure 3.10. Functional Detail Observed in the Retina.

The Oculomotor system controls the direction of gaze; part of it is automatic and part is consciously controlled. Spatio-temporal DoG feature enhancement is a consequence of the photoreceptor response and horizontal and bipolar cell connectivity.

Coarse-coding in the Signal Frequency Domain

We extend the use of coarse-coding to the signal frequency domain by considering Gaussian curves that simulate signal-processing filters. Gaussian-based filters were chosen due to the Gaussian nature of various stages of neuronal processing in vision systems as well as the ease of implementing Gaussian filters in electronic systems.

The Gaussian-based filters with different variances and their power spectra are shown in Figure 3.1–11. Gaussian curves *G1* through *G4* have increasing variances. Each curve is normalized so that the peak is at the same location. This way, the shape of the curve can be observed. In practical applications, the curves would be normalized for unity area so that filtering changes the signal without adding or taking away energy.

The spectrum of these Gaussian filters is Gaussian with decreasing variances. A curve with a small variance, such as *G1*, will pass low and medium frequency components and attenuate high ones, while one with a larger variance, such as *G4*, will only pass very low frequency components. Subtracting these filters gives us the *Difference-of-Gaussian* (DoG) filters shown. For the variances selected, DoG *G1–G2* serves as a high-pass filter, while the others serve more as band-pass filters.

Keep in mind that frequency here implies signal frequency. The signal could contain variations in spatially distributed energy (*spatial frequency)*, variations of intensity with time at a single location (*temporal frequency*, or variations in color with respect to either time or space (*chromatic frequency*).

Pairs of filters can be selected to decompose a signal into selected specified frequency components. For example, if it is desired to measure the strength of a signal at around 10% of the sampling frequency (horizontal axis in Figure 3.1–11), then the difference between gaussians *G3* and *G4* would be used to filter the signal. Due to linearity of the Fourier Transform, the spectral responses (middle plot in Figure 3.11) can be manipulated by addition or subtraction to get the desired spectral response of the filter (bottom plot). This simply translates to the same manipulation in the signal domain (top plot).

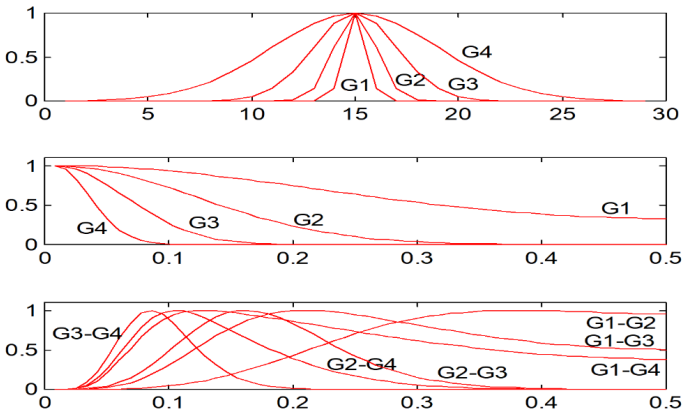


Figure 3.1–11. Selected Gaussian-based Filters.

All curves are normalized to a unity peak value. The top plot shows Gaussian filters plotted against data samples. The middle and lower plots show power spectra of Gaussian filters and DoG filters (respectively), plotted against sampling frequency.

Photoreceptor Mosaic

These filtering concepts are readily extended to two dimensions for use with the planar processing behavior of vision system models. To fully appreciate the nature of the image filter, it is essential to understand that the pixels are not uniformly distributed in size or type. The input image comes from a photoreceptor mosaic composed of *S, M,* and *L* cones and Rods.

Figure 3.1–12 shows a gross simplification of the photoreceptor mosaic. The central region is called the *fovea* and represents a circular projection of about a 1° conical view of the environment. In this region are only two photoreceptor types: *M* and *L* cells. Two cone types allow for color discrimination in the fovea, and the lack of rod cells allows for a high degree of spatial acuity. The rapid decline of spatial acuity with *eccentricity*, or the amount of separation from the center, can be clearly demonstrated by looking at a book on a bookshelf. Keeping the eyes fixed, it becomes difficult to read titles that are still relatively close to the fixation point.

The lack of rod cells in the fovea accounts for the disappearance of a faint star when we look directly at it. Rod cells are far more sensitive, so they respond in nighttime dim lighting conditions. However, if cones are not stimulated, there is no color discrimination since a strong signal at a frequency with weak response is the same as a weak signal at a frequency with strong response.

Figure 3.1–13 shows a representative mapping of fovea *L* and *M* cells into the parvo- (PP) and magnocellular (MP) pathways. The PP cells are physically smaller, but also carry information pertaining to smaller receptive fields. In the figure, the *L* and *M* ratios in the MP are kept nearly constant (2:1) so that the only response would be increased or decreased intensity (luminance). The PP surround cells, however, are skewed toward the cell not in the center. In other words, overall, there is a 2:1 ratio of *L:M* cells. The surround field in the upper left connection is 1:1, which favors the *M* cell contribution when the *L* cell is the center. The other example (upper right), the surround is purely *L*, which favors *L* over the 2:1 ratio when *M* is in the center. The surround, therefore, is at a slightly different cellular concentration that helps to favor local contrast between the two spectrally different cone types, allowing for a stronger acuity in the chromatic domain.

3.1.4 Color Vision Processing Models

There are several ways to designate the three cone types shown by their spectral responses in Figure 3.1–2. Some researchers use *B*, *G*, and *R* to represent *blue*, *green*, and *red* peaks in the photon absorption curves, although the peaks are not at those precise colors. Others prefer to use *S*, *M*, and *L* to denote the *short* wavelength, *medium* wavelength, and *long* wavelength responses, respectively. This latter designation is more appropriate since the notation in Boynton's model is changed to keep consistency between

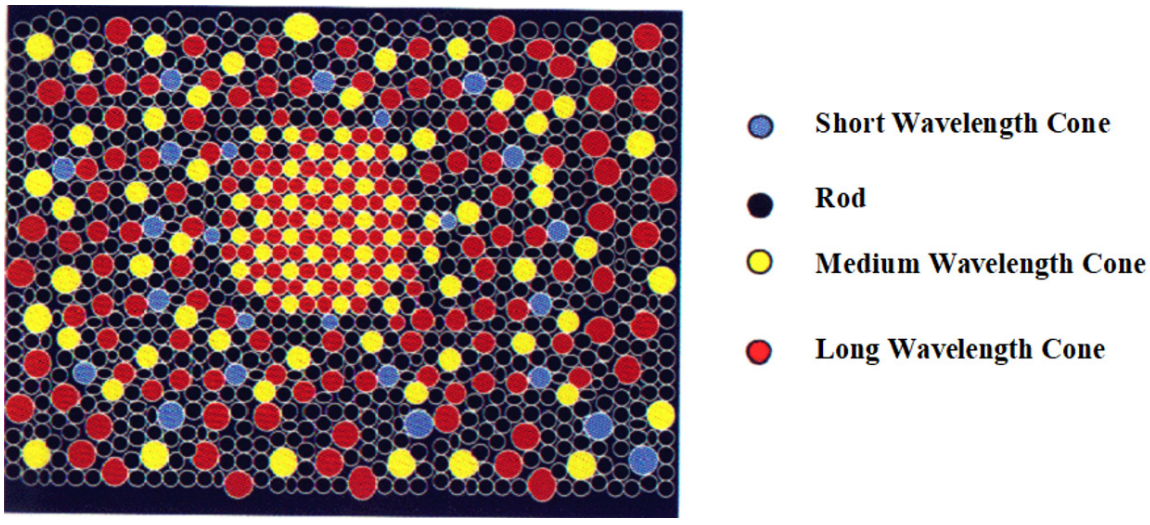


Figure 3.1–12. Photoreceptor Mosaic.

Three cone types and rod cells shown in representative distribution with eccentricity. *L:M* ratio is about 2:1 in the fovea with no rod cells, but overall rod cells far outnumber cone cells. Cone cell sizes increase slightly with eccentricity.

the three models presented in the next sections. All three describe separate luminance and chromatic channels of information within color vision processing.

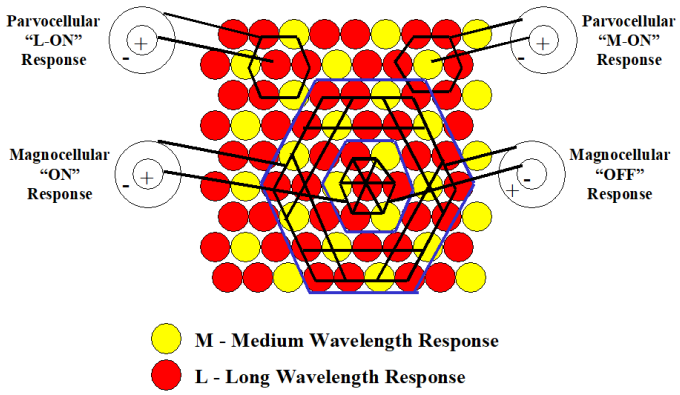


Figure 3.1–13. Representative Pathway Mapping.

"On" and "Off" imply polarity of the difference between the center and surround. The dark lines connecting cell centers represent signal connectivity in the area. Each area (center or surround) is represented by a single integrated value. MP cells contrast larger concentric receptive fields while the PP cells contrast individual cells with adjacent cells.

Guth Color Model [Guth91]

A model proposed by Guth included luminance and chromatic channels, as shown in Figure 3.1–14. The response of the luminance channel can be summarized as L+M, while the response of the chromatic channel can be described as L - S. A variation of this model mixes chromatic and luminance channels with automatic gain control in an artificial neural network trained by psychophysical data. The localized gain control simulates the spatial-temporal characteristics of the photoreceptor-horizontal cell network. There are numerous research efforts that have used various methods of emulating lateral inhibition for the spatial-temporal feature extraction inherent in the photoreceptor-horizontal cell network.

The first stage of the Guth model is the summation of simulated receptor noise sent to each cone followed by a steady-state self-adapting nonlinear gain control. The second stage is linear combinations of signals divided into two sets of three channels each. The third stage is a nonlinear compression of the second stage channels. One set includes two opponent channels and one non-opponent channel compressed to provide visual discriminations and apparent brightness. The other set includes three channels compressed to provide the appearances of light in terms of whiteness, redness or greenness, and blueness or yellowness [Guth91, Guth96].

This model has been criticized as being a poor emulation of retinal structure since no provision is made for cone proportions, the nature of anatomical connections, and the receptive field structure of ganglion and geniculate (LGN) neurons. Also, it appears to be an artificial neural network, with no physiological basis, which is trained to fit psychophysical data [DeV96]. Nevertheless, the division of color processing into luminance and color channels is an integral part of the model, and the point here is that several of these models include similar arrangements of cone types for these vision channels.

Boynton's Color Model [Boyn60]

A classic model by Boynton also divides the color vision pathways into luminance and chromatic channels. The luminance channel in his model is described as *L+M*. The chromatic channels are described as *L-M* and (*L+M*) - *S*. He points out the similarity in numerous others. The opponent chromatic channels are known from recordings at the *horizontal* cell layer. The horizontal cells connect to the photoreceptors and perform spatial and temporal photoreceptor signal mixing. The bipolar cells are thought to propagate difference signals in the opponent pathways [Boyn60].

DeValois' Color Model [DeV88]

A later model proposed by DeValois (Figure 3.1–14) goes into more detail by considering the relative concentrations of cells into account. It is observed that the concentration of the various cone cells is a function of *eccentricity*, or the distance from the center. In the center, the *foveola*, there are only *L* and *M* cells in a respective ratio of about 2:1. *S* cones become more apparent in the *parafovea* and more peripheral regions of the retina. There is an overall presumed ratio of *L:M:S* cells of 10:5:1. The normalized response of a neighborhood with these concentrations gives:

$$DeV\_LMS = 0.625L + 0.3125M + 0.0625S.$$

The variable *DeV_LMS* represents the response from a typical photoreceptor neighborhood with representative cell population densities. The DeValois color model consists of 4 center-antagonistic-surround channels, 3 representing PP channels and one representing an MP channel. Each of the 4 channels exists in two polarities for a total of 8 channels. The 6 chromatic channels model PP channel responses as

$$PP_L = (+/-)\ (L - DeV\_LMS)$$
$$PP_M = (+/-)\ (M - DeV\_LMS)$$
$$PP_S = (+/-)\ (S - DeV\_LMS)$$

while the luminance channels model the MP channel responses as

$$MP = (+/-)\ ((L + M) - DeV\_LMS)$$

The general concept for the Guth and DeValois color vision model is illustrated in Figure 3.14.

Generic Color-Opponent Model

The Boynton and DeValois models along with models from Martinez-Uriegas [Mart94] and Chittka [Chittka96] are compared in Figure 3.1–15. All of these (as well as Guth) have some sort of *L* and *M* cell synergism for encoding luminance and cell antagonism for encoding color. (*N* and *W* in Martinez-Uriegas model are for *narrow* and *wide* receptive field areas. *S* in the other models are for *small-wavelength* cones). Based on these popular models a simple color model could include a center receptive field contrasted with its local neighborhood. The center receptive field is modeled as a single picture element, or *pixel*. Ratios of the center pixel with the local neighborhood represent the color-opponent response. The models presented use differences, but ratios are in this generic model. This is plausible since many neurons respond logarithmically with stimulus, and ratios become differences after a logarithmic transformation. The actual responses of bipolar cells are presumed subtractive, but they can be considered divisive since the subtraction follows the logarithmic response of the photoreceptors.

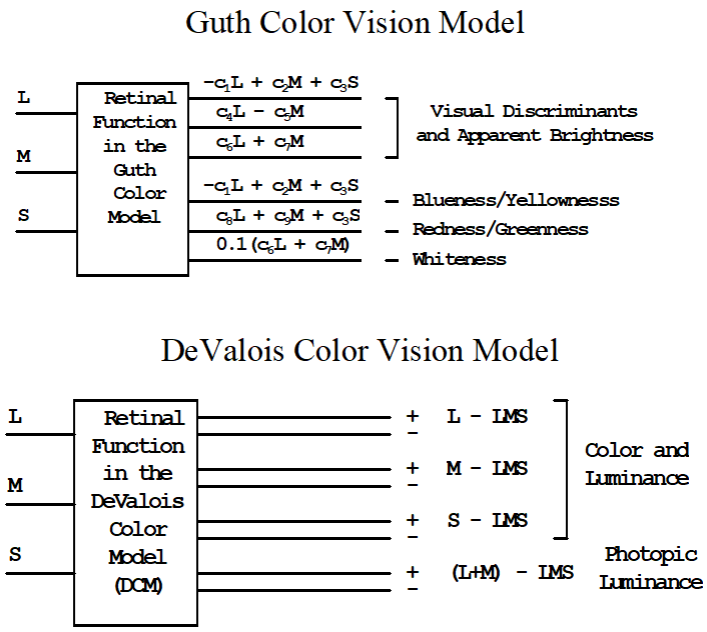


Figure 3.1–14. Guth and DeValois Color Vision Model.

These models and others generally include additive *L* and *M* response for luminance and subtractive *L, M,* or *S* response with the surrounding neighborhood for color.

	Chromatic Channel	Luminance Channel
Boynton	**L-M**	**L+M**
	(L+M)-S	
DeValois	**L-LMS**	**(L+M)-LMS**
{LMS = (10L+5M+S) / 16}	**M-LMS**	
	S-LMS	
Martinez-Uriegas	**(L-M)(N+W)**	**(L+M)(N-W)**
	(M-L)(N+W)	**(L+M)(W-N)**
Chittka	**S vs ML: a=1**	
{Excitation = aEs + bEm + cEl, a+b+c=1}	**M vs SL; b=1**	
	L vs SM: c=1	

Figure 3.1–15. Various Color Vision Model.

Common among these models is a local contrast between *L* and *M* responses for chromatic information and a broad integration of *L* and *M* responses for intensity information.

The photoreceptor responses are believed to be *logarithmic*, while the bipolar cell responses are believed to be *subtractive*. Due to the logarithmic nature of the photoreceptor response, the bipolar difference signal really reflects a contrast *ratio* of the photoreceptor with the horizontal-cell-mediated signal (which is a localized spatial-temporal average signal). This is because a logarithm transform of the ratio reduces a multiplication to an addition. For example, if an M detector responds with an output value of M_o and an L detector responds with an output value of L_o, then the logarithm of the ratio is the same as a subtraction of the individual logarithm-transformed cell responses. That is,

$$\ln (M_o / L_o) = \ln(M_o) - \ln(L_o).$$

3.1.5 Extracting color from parvocellular color-opponent pathway

Figure 3.1–13 shows on and off parvocellular pathways as a difference between a single photoreceptor cell in the center and a local neighborhood of a few adjacent photoreceptors. A representative photon absorption curve for each receptor (S, M, L, and Rod) is shown in Figure 3.1–2 If the neighboring receptors are averaged together the average response will be different form the center cell's response because on average the response of the center field is different from that of the neighborhood. To illustrate this concept, consider this example:

Example 3.1 Center-Surround Opponent Processing

Given photoreceptor spectral response curves in Figure 3.1–2 and a unity-intensity mono-chromatic stimulus determine the output of a center-surround antagonistic. Assume the surround input is made of a ratio of long-wavelength (L) to medium-wavelength (M) to short-wavelength (S) cones of L:M:S = 10:5:1. Assume the center field is only one cell (L, M, or S). Determine the output for a center cell of each cell type (S, M, and L) for a stimulus whose wavelength is

a) 450 nm
b) 500 nm
c) 550 nm
d) 600 nm

Solution:

Using Figure 3.1–2 we need to estimate the response of each stimulus that is expected from each of the three cell types. Looking at the normalized values at 450 nm the S-cone response is about 0.6, the M-cone about 0.3, and the L-cone about 0.1. The estimated measurements are shown in Figure 3.1–16. If the center cell is an S-cone cell the center value is 0.6. The surrounding neighborhood is calculated as a weighted average of the different responses. For L:M:S = 10:5:1 then the weighted average would be

$$\text{surround\_response} = \frac{1}{16}(10(0.1) + 5(0.3) + (0.6)) = \frac{3.1}{16} = 0.194$$

and the S-cell center-surround response would be

$$\text{S cell: center\_response} - \text{surround\_response} = 0.6 - 0.194 = 0.406$$

Similarly, at 450 nm,

$$\text{M cell: center\_response} - \text{surround\_response} = 0.3 - 0.194 = 0.106$$

$$\text{L cell: center\_response} - \text{surround\_response} = 0.1 - 0.194 = -0.094$$

Then the same can be done at 500, 550, and 600 nm. The following figure shows an estimated measured response for all three cell types at each of the 4 wavelengths:

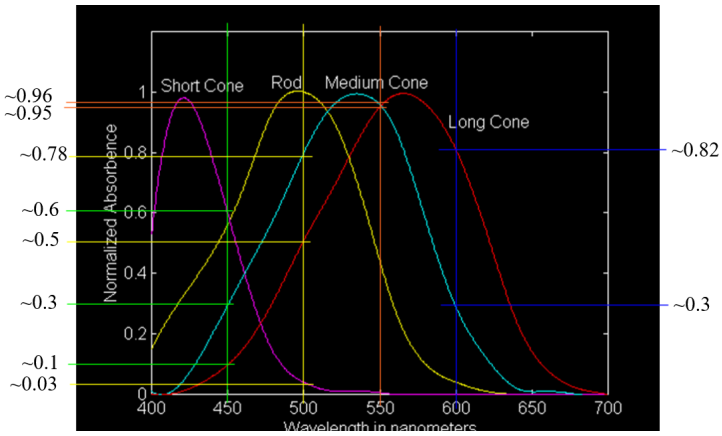


Figure 3.1–16 Estimated normalized absorption values for Example 3.1.

See Example 3.1

Using the weighted average as before, the result for each of the three cell types for each of the four wavelengths are:

Stimulus Wavelength	Center-surround opponent response		
	S-cell	M-Cell	L-cell
450 nm	0.41	0.11	-0.09
500 nm	-0.53	0.22	-0.06
550 nm	-0.89	0.07	0.06
600 nm	-0.61	-0.31	0.21

■

Looking at the results of this example we see positive responses in the forward diagonal and negative responses away from it. This makes sense as the input wavelengths used for this example are incrementally increasing as are the peak response wavelengths going from *S* to *M* to *L* cell. When the input stimulus is near the peak response of the center cell then the weighted average of the local neighborhood is lower since it is influenced by cells not responding as strongly. Of course, this contrast is far more significant in the PP channel than the MP channel since the PP center field is typically a single cell instead of an aggregate of cells in a typical MP channel. The contrast caused by color is therefore much stronger in the PP channel than the MP channel, which is why color is attributed to the PP channel in Figure 3.1–8.

This example assumes an object emitting (or reflecting) energy at a single monochromatic frequency, but most natural objects emit a wide distribution of frequencies across the visible spectrum. Regardless of the chromatic frequency distribution the algorithm results in a single specific response for each input that the higher brain processing can use to perceive a specific color. The color difference of an object against its background is amplified by this contrast, which benefits a species dependent on color perception for survival.

3.1.6 Gaussian Filters

One of the original models for the *outer plexiform* layer (photoreceptor-horizontal-bipolar cell interconnection layer) is the *Laplacian-of-Gaussian* (LoG) filter. For a gaussian function, *G,* defined in terms of a radius from the center, *r*, so that $r^2 = x^2 + y^2$ for cartesian coordinates *x* and *y*, then *G* is defined in terms of the variance, σ, as

$$G = e^{\frac{-(x^2+y^2)}{2\sigma^2}} = e^{\frac{-r^2}{2\sigma^2}}$$

Gaussian Filter

The LoG filter is defined as the second derivative of *G*:

$$\nabla^2 G(r) = \frac{-1}{\pi\sigma^4}\left(1 - \frac{r^2}{2\sigma^2}\right)e^{\frac{-r^2}{2\sigma^2}}$$

Laplacian-of-Gaussian (LoG) Filter

The *Difference-of-Gaussian* (DoG) for two gaussians with variances σ_1 and σ_2, is

$$G_1 - G_2 = e^{\frac{-r^2}{2\sigma_1^2}} - e^{\frac{-r^2}{2\sigma_2^2}}$$

Difference-of-Gaussian (DoG) Filter

Under certain conditions, the DoG filter can very closely match the LoG filter [Marr82]. The DoG filter allows more flexibility as two variances can be modified, thus there are two degrees of freedom. The LoG filter only uses one variance, thus only one degree of freedom.

The spectrum of a gaussian is also a gaussian:

$$e^{-t^2/2\sigma^2} \Leftrightarrow \sigma\sqrt{2\pi}e^{-\sigma^2\omega^2/2}$$

Note that the variance, σ^2, is in the denominator of the exponent in the time domain and in the numerator of the exponent in the frequency domain. This is shown graphically in Figure 3.1–11 as the broad (large variance) gaussians result in sharp spectral responses, passing only very low frequencies. The narrow (small variance) gaussians pass more of the lower and middle frequencies. The limits are a zero-variance gaussian, which, when normalized to unity area, becomes the impulse function, and an infinite-variance gaussian, which becomes a constant. An impulse function passes all frequencies, and a constant only passes the DC component of the signal, which, in frequency domain, is represented as an impulse at $\omega = 0$ (repeated every 2π increment of ω due to the periodicity of the Fourier Transform:

$\delta(t) \Leftrightarrow 1$ Zero-variance gaussian limit

$1 \Leftrightarrow 2\pi\delta(t)$ Infinite-variance gaussian limit

3.1.7 Wavelet Filter Banks and Vision Pathways

The two primary vision pathways are the magnocellular pathway (MP) and the parvocellular pathway (PP). Each neuronal response in the MP represents a local average over a large receptive field. Each neuronal response in the PP represents local detail in a smaller receptive field. Thus,

the MP and PP decompose the natural input image into local average and local detail components, respectively.

Similarly, digital images can also be decomposed into a set of averages and another set of details using *quadrature mirror filtering* (QMF). This method of image analysis (breaking apart images into components) and synthesis (reconstructing images from the components) results in a series of averaging components and another series of detailing components [Strang96]. QMF is a special case of *sub-band coding*, where filtered components represent the lower and upper frequency halves of the original signal bandwidth. If the analyzing filter coefficients are symmetric, then the synthesizing components are mirrored with respect to the half-band value, thus the term *quadrature mirror*. The structure of such a wavelet analyzer and synthesizer is shown in Figure 3.1–17. The low pass filter (LPF) and high pass filter (HPF) are similar in functionality to the MP and PP in time, space, and color domains. A variety of applications have emerged from the QMF.

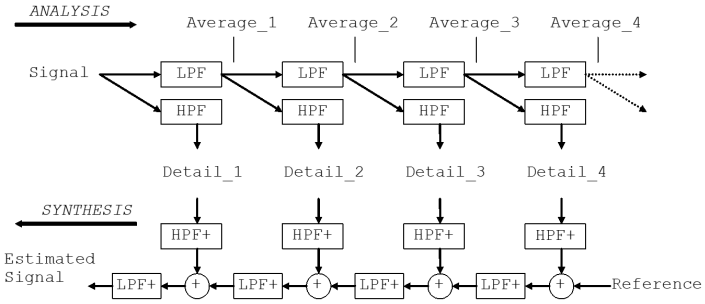


Figure 3.1–17 Quadrature Mirror Filtering (QMF).

QMF requires low-pass and high-pass filters whose spectra cross at one-fourth the sampling frequency and mirror each other in spectral response. Each LPF and HPF in the analysis path includes a down-sampler, while each adjoint LPF and HPF in the synthesis path includes an up-sampler (or interpolator).

To illustrate QMF the following example and exercise decomposes a sequence into its averages (after LPF) and details (after HPF). The sequence is down-sampled after each pass through the LPF; all LPFs are the same and all HPFs are the same (technically, the reconstruction filters are adjoint filters, but are the same for real-valued coefficients).

Example 3.2, 1D QMF Analysis and Synthesis

a) Using the discrete Harr wavelets [0.5 0.5] and [0.5 -0.5] for LPF and HPF respectively, show how to decompose the following sequence into one average value and a set of detailed values.

b) Reconstruct the original sequence from the calculated components to verify correct decomposition.

c) Compare the energy of the original sequence with the energy of the components.

$$x[n] = \{12 \;\; 16 \;\; 8 \;\; 10 \;\; 10 \;\; 18 \;\; 13 \;\; 17\}$$

Solution:

Figure 3.1–18 shows the QMF symmetry of the PSD for the given LPF and HPF.

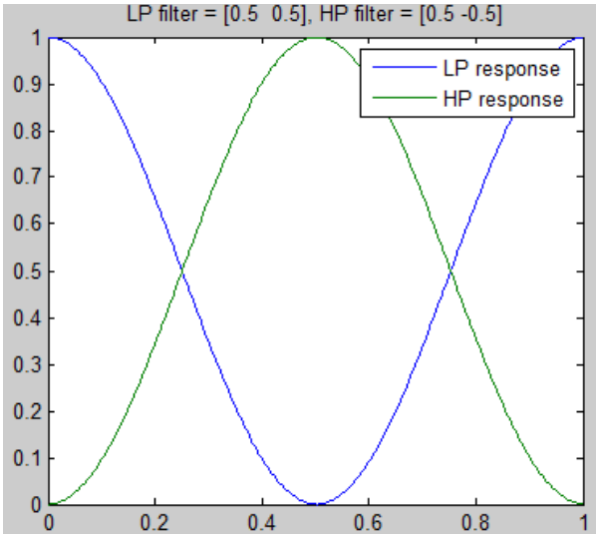


Figure 3.1–18 Power Spectral Density of Harr LPF and HPF showing QMF symmetry.

Horizontal axis is sampling frequency. See Example 3.2

Part a:

We now filter the input sequence with the LPF and HPF (and stop once we have the same number of values, thus discarding the last value). Using the graphical method of convolution, flipping the LPF (which is symmetrical) and passing under $x[n]$, taking dot product, and shifting results in

$x[n]$:		12	16	8	10	10	18	13	17	
LPF[-n]:	0.5	0.5								= 6
		0.5	0.5							= 14
			0.5	0.5						= 12
				0.5	0.5					= 9
					0.5	0.5				= 10
						0.5	0.5			= 14
							0.5	0.5		= 15.5
								0.5	0.5	= 15

First LPF result is {6 14 12 9 10 14 15.5 15}
Down-sampling LPF results gives {14 9 14 15}, which will be the input to the next LPF stage.

Similarly, using the graphical method of convolution, flipping the HPF and passing under $x[n]$, taking dot product, and shifting results in

$x[n]$:		12	16	8	10	10	18	13	17	
HPF[$-n$]:	-0.5	0.5								= 6
		-0.5	0.5							= 2
			-0.5	0.5						= -4
				-0.5	0.5					= 1
					-0.5	0.5				= 0
						-0.5	0.5			= 4
							-0.5	0.5		= -2.5
								-0.5	0.5	= 2

First HPF result is {6 2 -4 1 0 4 -2.5 2}
Down-sampling HPF results gives {2 1 4 2}, which will be saved as detailed components.

To determine the results of the second stage we repeat the LPF and HPF on the down-sampled LPF results of the first stage:

Down-sampled first-stage LPF results:		14	9	14	15	
LPF[$-n$]:	0.5	0.5				= 7
		0.5	0.5			= 11.5
			0.5	0.5		= 11.5
				0.5	0.5	= 14.5

Second LPF result is {7 11.5 11.5 14.5}
Down-sampling gives {11.5 14.5}, which will be the input to the next LPF stage.

Down-sampled first-stage LPF results:		14	9	14	15	
HPF[$-n$]:	-0.5	0.5				= 7
		-0.5	0.5			= -2.5
			-0.5	0.5		= 2.5
				-0.5	0.5	= 0.5

Second HPF result is {7 -2.5 2.5 0.5}
Down-sampling gives {-2.5 0.5}, which will be saved as detailed components

To determine the results of the third stage we repeat the LPF and HPF on the down-sampled LPF results of the second stage. Subsequent down-sampling results in one value with will be saved:

Down-sampled second-stage LPF results:			11.5	14.5
LPF[-*n*]:	0.5	0.5		= 5.75
		0.5	0.5	= 13

Third LPF result is {5.75 13}
Down-sampling gives the value 13. This value represents the sequence average.

Down-sampled second-stage LPF results:			11.5	14.5
HPF[-*n*]:	-0.5	0.5		= 5.75
		-0.5	0.5	= 1.5

Third HPF result is {5.75 1.5}
Down-sampling gives the value 1.5, and the analysis is complete.

A summary of the filter outputs is listed here, and the value after down-sampling is underlined:

First LPF result:	{6 <u>14</u> 12 <u>9</u> 10 <u>14</u> 15.5 <u>15</u>}
First HPF result:	{6 <u>2</u> -4 <u>1</u> 0 <u>4</u> -2.5 <u>2</u>}
Second LPF result:	{7 <u>11.5</u> 11.5 <u>14.5</u>}
Second HPF result:	{7 <u>-2.5</u> 2.5 <u>0.5</u>}
Third LPF result:	{5.75 <u>13</u>}
Third HPF result:	{5.75 <u>1.5</u>}

The QMF components in $x[n]$ are the down-sampled HPF results and the final average, which is the sequence {2 1 4 2 -2.5 0.5 1.5 13},
where the last value is the sequence average.

Part b:

For the purposes of this text, which is to illustrate reconstruction from the components, we will simply subtract the detail from the average and then add the detail to the average to show the original sequence can be reconstructed. The final detail, 1.5 will be subtracted from the final average, 13, to give 11.5, and then the same two values will be added to give 14.5:

Reconstructing second stage: { (13-1.5) (13+1.5)}
= {11.5 14.5}

Then the second-stage down-sampled detail, the sequence {-2.5 0.5} will be used to subtract and add to the second stage average values just determined above:

Reconstructing first stage:

{(11.5-(-2.5)) (11.5+(-2.5)) (14.5-0.5) (14.5+0.5)}
= {14 9 14 15 }

and the original sequence determined from those value minus then plus the down-sampled first-stage details:

$x[n]$ = {14-2 14+2 9-1 9+1 14-4 14+4 15-2 15+2}
= {12 16 8 10 10 18 13 17}

Part c:

One of the benefits of decomposition is the great reduction in signal energy. The total energy is the sum of the square of each of the components, which results in

Power in $x[n]$:
$12^2 + 16^2 + 8^2 + 10^2 + 10^2 + 18^2 + 13^2 + 17^2 = 1446$

Energy in QMF components of $x[n]$:
$2^2 + 1^2 + 4^2 + 2^2 + (-2.5)^2 + 0.5^2 + 1.5^2 + 13^2 = 202.8$

As sequences become larger and signals become multidimensional (such as images or image sequences) the comparison can be far more dramatic (orders of magnitude).

■

Exercise 3.1, 1D QMF Analysis and Synthesis

Using the discrete Harr wavelets [0.5 0.5] and [0.5 -0.5] for LPF and HPF respectively, show how to decompose the following sequence into one average value and a set of detailed values.

$$x[n] = \{2\ 22\ 4\ 12\ 0\ 16\ 0\ 4\}$$

Answer:
QMF Components of $x[n]$: {10 4 8 2 -2 -3 -2.5 7.5}, where the last value is the sequence average.

■

Vision pathways (MP and PP) and QMF filter banks both therefore break up the input image signal into high and low frequency components. The MP and PP are further augmented by the rod-system pathway. Rod cells are highly interconnected and although the rods themselves are basically saturated in daylight conditions; the rod bipolar cells are mediated by neighboring cone cells. The overall effect is a spatial low-pass filter of the mosaic image.

A model of the low frequency rod system filter can be combined with a model of the PP to create a pair of filters whose spectral response crosses at one-fourth the sampling frequency, or half the Nyquist-limited frequency. A carefully chosen pair can give a striking resemblance to typical filter pairs chosen for QMF applications. A model of the MP can be substituted for the low frequency filter, but the spectral response will diminish with very low frequencies.

3.1.8 Coarse Coding and the Efficient Use of Basis Functions

Natural vision systems process information in space, time, and color domains. In each of these domains we find filters that are typically few and relatively coarse in bandwidth. There are essentially only four chromatic detector types, three temporal channels, and three spatial channels. The responses of these elements must be broad in scope to cover their portion of the data space. For example, in daytime conditions only three detector types have varying responses. As a minimum each type must cover one-third of the visible spectrum.

Coarse coding resembles the more common wavelet applications typified by complementary coarse low pass and high pass filters. QMF signal reconstruction capability is a practical demonstration of extracting specific spectral detail from only two broadband filters. An interesting corollary to this line of research is that the behavior of such synthetic applications may lead to a deeper understanding of natural information processing phenomena.

3.1.9 Nonorthogonality and Noncompleteness in Vision Processing

Sets of wavelets can be subdivided into *orthogonal* or *nonorthogonal* and *complete* or *noncomplete* categories. A set of functions is *orthogonal* if the inner product of any two different functions is zero, and *complete* if no nonzero function in the space is orthogonal to every vector in the set. Orthogonality provides computational convenience for signal analysis and synthesis applications. Completeness ensures the existence of a series representation of each function within the given space. Orthogonality and completeness are desired properties for wavelet bases in compression applications.

However, biological systems are not concerned with information storage for perfect reconstruction. Any machine-vision application requiring some action to be taken based on an understanding of the image content will also fit this general description. In fact, many biological processes can be modeled by sets of functions that are nonorthogonal [Daug88]. The task is processing information to take some action, not processing information for later reconstruction. Using nonorthogonal filters leads to a redundancy of information to cover the span of information. The redundancy of vision filters is balanced by the need for efficiency, simplicity, and robustness. Information redundancy results in unnecessary hardware and interconnections, but often redundancy may be required to sufficiently span the information space inherent in the

environment. The cost of supporting the redundancy may be less significant than the benefit of using simpler processing elements that degrade gracefully. Since there is a closeness between Gaussian-based filters and more mathematically elegant filters (such as Laplacian) there is good retention of pertinent information (though not perfect).

3.2 Applications inspired by natural photo-sensory Systems

The first photosensory application is the author's own idea to use gaussian filters for emulation of low-pass spatial-temporal filters of the photoreceptors and horizontal cells and to do that at three levels, each resulting in inherent delays that are used for elementary motion detection (EMD) models. The three different levels allow for modeling the well-known center-surround contrasting signals (propagated by bipolar cells) that comprise the magnocellular and parvocellular pathway signals. It also allows for two different EMD's at each location. The additional EMD gives a degree of freedom needed to determine edge velocity.

The next group of research efforts are focused on modeling the outer plexiform layer (OPL) of the retina (photoreceptors, horizonal cells, and bipolar cells) using VLSI circuits. Biology is made of material with a natural plasticity for adapting to the organisms needs. Silicon is brittle, but very reliable as a technology for implementing the behavior of the OPL. Following those efforts are the ones combining the silicon retina concepts with optic flow for a more comprehensive adaptive pixel that better emulates the OPL.

A few examples of exploiting natural foveal vision are then presented. The densely-packed photoreceptors in the very center of the retina provides much better spatial acuity than the periphery, where photoreceptors are not as densely packed. This can be misleading as our ability to see detail in the very center far surpasses that of the periphery, and the photoreceptor packing is a very small part of that. There are about 5 times the number of rod cell than cone cells in the retina, but none in the fovea (thus a faint star may disappear when we look right at it). Also, cells are more interconnected in the periphery to afford better temporal resolution at the cost of spatial resolution. Most fovea-vision-inspired applications concern the higher resolution in a region of interest and not the representative rod and cone cell distributions and non-uniform level of cell interconnections.

The group following is focused on asynchronous event-based signaling which, like biology, results in a spike (or action potential) when a significant event happens (or a threshold is exceeded). Diverging from biology into a possible realm of much higher signal processing capabilities is the notion of doing the same OPL signal processing but with photonics rather than electronics. This would be a significant deviation from biology, but as pointed out before many times researchers are using biology to glean novel ideas and not necessarily attempting to duplicate biology. Another frontier being pursued is the incorporation of polarization information in vision systems as indicated in the final section.

3.2.1 Combined EMD and Magno/Parvo channel model [Brooks18]

The Hassenstein-Reichardt elementary motion detection (HR-EMD) model [Hass56] reviewed earlier cannot accurately measure optic flow velocity. A simplified version of the HR-EMD is shown in Figure 3.2.1–1. There is an optimal speed for the peak response of the EMD based on the design of the delay element. If the spatial contrast is weak but moving across the image at that speed the response can be moderate and can be the same as a stronger spatial contrast moving at a sub-optimal speed. Another information dimension is needed to determine edge velocity; one approach is to measure the power spectral density (PSD) of the image and combine that with a global EMD response in the form of a look-up table [Wu12] although a PSD measurement of an image is not known to exist in biology. These and similar approaches have used the delay inherent in traditional low-pass filters (LPFs) such as Butterworth filters (popular due to being maximally flat in pass band). Again, Butterworth filters are not known in biology. The best model of LPFs in biology are gaussian filters, which are not popular in conventional applications due to properties such as non-orthogonality. However, gaussian filters are naturally occurring in biology due to ion leakage, charge-sharing amongst receptor cells and excitatory and inhibitory signals of adjacent layers of neurons.

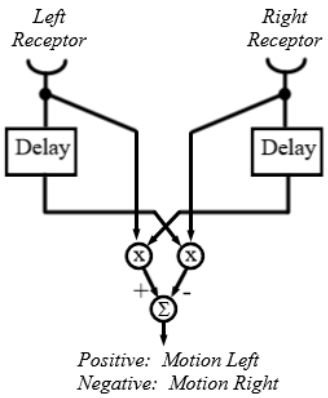


Figure 3.2.1–1. Basic HR-EMD correlation model.

A moving object stimulating one receptor then the other will trigger correct motion on both leading and trailing edges of object stimulation; the object intensity can be greater than or less than the background.

Gaussian filters can also model the magnocellular and parvocellular pathways (MP and PP); each channel of the MP or PP can be modeled as a difference-of-gaussian filter between the center receptor (or group of receptors) and the surrounding receptors, referred to as center-surround antagonistic signals. To model the either the MP or PP two gaussian are needed, a smaller variance gaussian for the center field and a larger variance gaussian for the surrounding field. Possibly (a subject for future experimentation) both channels can be modeled with a total of 3 gaussian filters, where the variance of the surrounding PP signal is the same as the variance of the center MP signal. These three gaussian filters are identified in Figure 3.2.1–2 as having high, medium, and low cutoff frequencies. Keep in mind these are *spatial-temporal* filters, so the frequencies are multidimensional to include both time and space. In the primate vision system these spatial-temporal filters would be implemented at each receptor location by the effects of weak inter-photoreceptor connections, the effects of lateral inhibition of the horizontal cells, the propagation of bipolar cells, the further mediation by the amacrine cells as the signal is passed through the ganglion cells.

Spatial-temporal gaussian filter effects are well known in vision. The three gaussians in Figure 3.2.1–2 provide the necessary information for both MP and PP channel modeling as well as two separate EMD channels, referred to in the figure as the *Parvo EMD* and *Magno EMD*. Having two separate EMD channels gives the additional degree-of-freedom needed for object velocity determination. The initial LPF (with high cutoff frequency) is used as the 'receptor' signal in Figure 3.2.1–1 for both EMDs, and the delayed signal is the output of the second LPF (medium cutoff) for the Parvo EMD while the delayed signal is the output of the third LPF (low cutoff) for the Magno EMD.

The object velocity is a function of location in the image, and ambiguity would be expected if only one EMD measurement were available. However, in this model two independent EMD outputs are available, so the object velocity would be determined by some combination of the responses of the Parvo EMD and Magno EMD. Another subject for future experimentation would be how the signals are combined to give the unique velocity. This is very consistent with the coarse coding concepts we see throughout biological sensory systems (and likely higher brain function).

Figure 3.2.1–3 shows how the two separate EMDs can be combined to give a specific object motion velocity at the given location in the receptive field. The output of the *left* and *right* receptors in this figure would be the output of the *high* cutoff LPF of Figure 3.2.1–2. The output of delays D_1 and D_2 correspond to the outputs of the *medium* cutoff LPF and the *low* cutoff LPF of Figure 3.2.1–2, respectively.

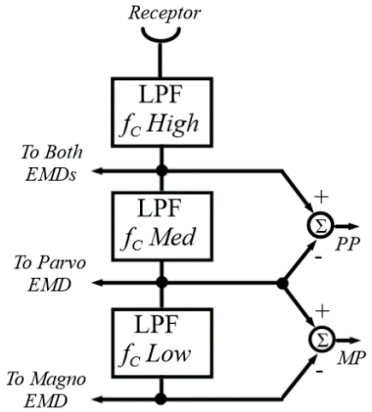


Figure 3.2.1–2. A 3-gaussian model for two EMD models and MP and PP models.

Each pixel represents a sensory receptor in 2D space. As the pixel update algorithm is repeated, the spatio-temporal cutoff frequency of the LPF is reduced. The outputs of three sets serve to provide biologically-relevant PP and MP signals as well as inputs to EMDs that can provide very accurate object velocity and direction. The use of *Parvo* and *Magno* EMD is used to distinguish which output set is used, as the same signal is used for the negative input to the respective channels.

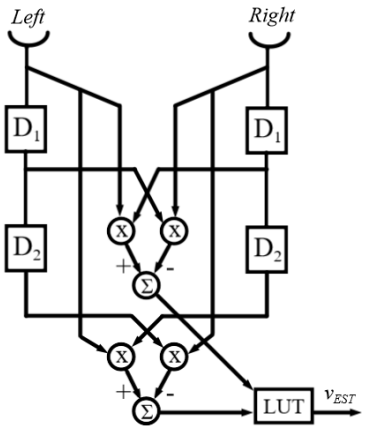


Figure 3.2.1–3. Combining two EMDs for accurate object velocity estimation.

The second and third set of pixel update iterations show in Figure 3 is shown here as delay elements D_1 and D_2 respectively. The input to the lower (or Magno) EMD allows the combination of delays for the two sets to be used. Because the delays are different, each EMD will be tuned to a different value in the range of motion. The look-up table (LUT) could be a simple comparison to determine a precise velocity from the two samples provided by the individual EMDs.

The effectiveness of this magno/parvo EMD model can be simulated in MATLAB or other visualization tool. Letting γ control the amount of spatial spreading between frames (limiting it to a value between 0 and 1) then the pixel value retained will be γ times the current value which will be added to (1- γ) times the average of the 4 nearest neighbor current pixel values. Letting α control the amount of temporal smoothing so that the current pixel value is multiplied by α and added to (1- α) times the current spatially-processed pixel. The spatial-temporal effects are provided by the horizonal cells, so we reference that signal as $H_{i,j}$, where i is the row index and j is the column index. Letting $P_{i,j}$ represent the pixel value (the modeled receptor value) at the *ith* row and *jth* column and using T as a temporary variable (for clarity) we have the following update algorithm:

$$T = 0.25(1- \gamma)(H_{(i-1),j} + H_{i,(j-1)} + H_{i,(j+1)} + H_{(i+1),j}) + \gamma P_{i,j}$$

$$H_{i,j} = (1- \alpha)T + \alpha H_{i,j}$$

The constants γ and α represent levels of spatial smoothing and temporal smoothing respectively, which in both cases gives the low-pass filtering effects of a gaussian filter (simultaneously in both time and space domains). These can be made adaptive once a performance metric is determined. Simulating the three filters in Figure 3.2.1–2 is accomplished by tapping the results at differing numbers of iterations as the visual information is processed. A few iterations implement a high cutoff (spatial-temporal) frequency, more iterations would give a medium cutoff frequency, and even more iterations a lower cutoff frequency. There are several degrees-of-freedom for experimentation, including the spatial and temporal smoothing constants along with the number of iterations for implementing the gaussian filters.

3.2.2 Autonomous hovercraft using insect-based optic flow [Roub12]

It is well known that insects such as honeybees navigate their environment by optic flow queues in the visual field. Insect-inspired optic flow was demonstrated in a small hovercraft robot [Roub12] autonomously following a wall and navigating a tapered corridor. The design is focused on obstacle avoidance in the azimuth plane with 4 2-pixel optic flow (OF) sensors at 45° and 90° on both left and right sides. The hovercraft robot followed a wall at a given distance as well as successfully navigating through a tapered corridor. As seen in experiments with honeybees [Srini11] the velocity decreases as it successfully navigates through a tapered corridor, a natural consequence of maintaining constant OF as side get closer. The honeybee navigation was presumed to be the result of the balancing OF on both sides of the insect.

The hovercraft demonstrated the ability to adjust forward speed and clearance from the walls without rangefinders or tachometers. However, a magnetic compass and accelerometer were used to prevent movement in the yaw axis direction so that the craft continues to move forward. This is necessary since the experiment focused on the OF queues and the ability to navigate the corridor.

The algorithm was developed in simulation and implemented on this hovercraft. All 4 sensors (two at 45° and two at 90° rom the forward direction on each side) were used in the navigation algorithm the authors call *dual lateral optic flow regulation principle*. It demonstrates a more comprehensive suggestion as to how honeybees navigate their environment than simply balancing optic flow from the two sides. This is an example of a bio-inspired sensor that is used to help biologists better understand how honeybees navigate their environment.

3.2.3 Autonomous hovercraft using optic flow for landing [Dup18]

In this effort 12 optic flow pixel sensors implementing a threshold-based motion detection is compared to a more traditional set of 12 optic flow pixels implementing a cross-correlation method. The cross-correlation method is more robust, but also more computationally complex. If a sufficient threshold method can work, then the complexity is greatly reduced. The drawback is the performance is strongly dependent on the threshold, which can vary from scene to scene and differing illumination conditions.

The application in mind is a hovercraft using optic flow sensing on the ventral side (under side) of the craft to ensure smooth landing. As an insect gets closer to the landing point the optic flow underneath will increase since the image texture is getting closer. If the insect keeps the optic flow constant, then its speed must be reducing as the insect approaches, until the point where the insect is at rest on the landing surface. To measure the performance the optic flow sensor was fixed with a textured visual field passed in front of the sensor.

3.2.4 Silicon Retina [Maha89]

The silicon retina [Maha89] is designed to emulate the initial processing layers of the retina, which include the photoreceptors, horizontal cells, and bipolar cells. An array of 48 x 48 pixels was fabricated using 2.0 µm design rules (width of conducting path) and pixel circuits about 109 x 97 µm in size. A hexagonal resistive grid is used so that local averages of pixels are more highly influenced by the six nearest neighbors than those farther away.

The triad synapse (connecting these three cell types) is modeled in silicon as a follower-connected transconductance amplifier. A capacitor stores the spatial-temporal signal of the photoreceptor, and an amplifier propagates the difference if this signal and the photoreceptor signal, modeling the bipolar cell center-surround antagonistic signal. The photodetector circuit is a bipolar transistor biased with a depletion region responding logarithmically with the incoming light intensity, which corresponds to physiological recordings of natural photoreceptors.

The design was later revised with an adaptive photoreceptor circuit modulated by three feedback paths and individual time constants. The gain of the receptor is adaptive, and the circuit was more robust to transistor mismatches and temperature drifts than the original silicon retina. Another improvement was the incorporation of the edge signal position without the need for off-chip subtraction [Maha91].

3.2.5 Neuromorphic IR analog retina processor [Mass93]

Building on the silicon retina design the Air Force Research Lab (AFRL, Eglin AFB) funded the development of an infrared sensor. One of the problems emulating biological retinae with VLSI technology is the area required to model the time constants observed in biology make the design of a 2D array of pixels unreasonably large. This IR sensor design used switch-capacitor technology with small capacitors to emulate time constants of larger capacitors. Although such technology has no biological counterpart, it was successful in achieving biomimetic spatial-temporal response rates. The drawback of this technology is additional noise caused by the 10KHz switching speeds required for the design.

A 128 x 128 array of Indium Antimonide (InSb) detector elements at 50 µm pitch were connected 4-to-1 to create a 64 x 64 array at 100 µm pitch. This detector plane was bonded to a readout chip where each pixel used the 100 µm pitch area for the switched-capacitor and readout circuitry. The InSb diodes were connected in photovoltaic mode and responded logarithmically as the biological photoreceptors do. The CMOS transistors configured as switched-capacitors were used between pixel nodes to provide the spatial-temporal smoothing inherent in laterally-connected horizontal cell layers of the retina.

The result was a medium-wave IR (MWIR) camera with localized gain control. The camera captured imagery of a gas torch in front of a lamp with a large flood light bulb. Conventional cameras at that time would saturate in all the lighted areas unless a global gain control were in place, in which case the objects in the darker parts of the image would not be seen. In this experiment the filament of the light bulb, the outline of the torch flame, as well as the object in the darker parts of the image could be clearly seen. This is the benefit of localized gain control of natural biological retinae and bio-inspired sensors that model them.

3.2.6 Michaelis-Menten auto-adaptive pixels M²APix [Maf15]

The vision system of primates (and other animals) provides responses over a wide range of luminosities while at the same time provides good sensitivity to local contrast changes, giving the vision system the ability to simultaneously distinguish a bright object against a bright background in one part of the image and a dark object against a dark background in another part of the image. The wide range of luminosities is facilitated by the open-

ing and closing of the iris as well as the natural logarithmic response of the photoreceptors. The good sensitivity is facilitated by the lateral inhibition of the post-photoreceptor processing neurons, the horizontal cells.

Many machine vision designers have sought to develop wide dynamic range sensors and have looked to the natural vision system for inspiration. The Delbruck adaptive pixel [Del94] used the logarithmic photoreceptor circuit of the original silicon retina [Maha88] and is used in comparison with the Michaelis-Menten auto-adaptive pixel (M^2APix) proposed here [Maf15].

The Michaelis-Menten equation [Mich1913] was derived to model enzyme kinetics in biochemistry. It describes the rate of enzymatic reactions in terms of the maximum rate achieved when the substrate is saturated and a constant representing the substrate concentration when the reaction rate is half the maximum rate [WikiMM]. It is adapted in [Maf15] to describe the photoreceptor's response, V, in terms of the maximum response at lamination saturation, V_m, the light intensity, I, and an adaptation parameter, σ, given in [Maf15] as

$$V = V_m \frac{I^n}{I^n + \sigma^n}$$

Substituting V with the enzymatic reaction rate, V_m with the maximum rate when the substrate concentration is saturated, I with the substrate concentration, and σ with the Michaelis constant, which is the substrate concentration when the rate is half V_m, and letting $n = 1$ this equation reduces to the original biochemistry equation [WikiMM].

The Delbruck adaptive pixel provides a 7-decade range of light adaptation and a 1-decade range of contrast sensitivity. There were some issues raised concerning steady-state responses increasing with light intensity and inconsistent transient responses under large contrast sensitivity. Other methods using resistive grids to emulate horizontal cell networks resulted in 4 decades of sensitivity but required external voltage sources to set bias points [Maf15].

A photoreceptor array of 12 M^2APix pixels and 12 Delbruck pixels was fabricated and used for comparison. The 2 x 2 mm silicon retina was fabricated into a 9 x 9 mm package with the two 12-pixel arrays side-by-side for comparison. The experimental results confirmed that the M^2APix pixels responded to a 7-decade range of luminosities and with a 2-decade range of contrast sensitivities. The advantage over the Delbruck adaptive pixel is that it produces a more steady contrast response over the 7 decades of luminosities so that the least significant bit (LSB) will be a lower value and therefore a better contrast resolution [Maf15].

3.2.7 Autonomous hovercraft using insect-based optic flow [Van17]

A bio-inspired eye is designed to allow an aerial vehicle passive navigation through corridors and smooth landing by having vision sensors responding to optic flow in the front, two sides, and the bottom. A given example would be a quadrotor exploring a building by keeping a certain distance from the walls.

Called the "OctoM^2APix" the 50 gm sensor includes 8 Michaelis-Menten auto-adaptive pixels (M^2APix): 3 measuring optic flow (OF) on the left side, 3 measuring OF on the right, and 2 on the bottom measuring OF on the ground underneath the vehicle. The center pixel on each side is measuring OF at right angles to the heading; one is pointing between the side and the front, while the other is pointing between the side and the rear. Each side covers about 92° in the horizontal plane. The object is to allow the vehicle to correct for heading based on the differing OF measurements of the three pixels on either side.

The experimental results were with the OctoM^2APix sensor stationary and a textured surface moving next to it at various angles with respect to the (simulated) heading, or the direction of the front of the sensor. The experimental heading included 0°, where the vehicle would be following a wall-like surface, +20°, where the (non-simulated) vehicle would eventually collide with the surface if not corrected, -20°, where the vehicle would be separating from the surface, and -45°, where the vehicle would be separating at a faster rate. The OF on forward and rear side pixels should offset each other when heading is parallel to the surface, and the difference between forward and rear side pixels would provide cues for the heading with respect to the wall, and thus allow the vehicle to adjust its heading if the goal were to follow the wall at a constant rate. The experimental results are shown by calculating heading from the center and forward pixel, the center and rear pixel, and the forward and rear pixel, the latter being the best estimate when all three sensors had the surface in view. This makes sense since this would be the widest separation.

3.2.8 Emulating fovea with multiple regions of interest [Azev19]

There are many applications where image resolution is high in some region of interest (ROI) and low in the remaining portion of the image. This is a crude resemblance of a foveated image but could be argued as bio-inspired by the fovea. In both natural and synthetic designs, the idea is to conserve computational resources by using a higher sampling in an ROI (center gaze for biology) and

lower sampling elsewhere. Non-uniform sampling is seen in all natural sensory systems as biology has adapted to the different levels of relevance of natural stimuli (passive or active). In many commercial and military applications multiple ROI's could be employed, but this is rare in biology if it exists. (Vision systems have a single fovea, but it could be argued that there are multiple regions of higher sampling, for example, in the sense of touch as the well-known somatotopic map would suggest).

A vehicle tracking system designed for self-driving cars uses multiple ROI's and claims fovea inspiration. The subsequent image processing is developed using deep-learning neural networks, which again implies some level of bio-inspiration. The system uses vehicle waypoints, which are the expected future locations of the vehicles, and continually crops the image looking for other vehicles. This is analogous to drivers looking down the road they are traveling. The experimental results claimed an improvement of long-range car detection from 29.51% to 63.15% overusing a single whole image [Azev19]. As pointed out before, many researchers are more focused on solving engineering problems (as they should be) and not too concerned with the level of biomimicry. Therefore, there can be a chasm between the levels of biomimicry between various efforts claiming bio-inspired designs.

3.2.9 Using biological nonuniform sampling for better virtual realization [Lee18]

There are other applications that are not mimicking biology but considering the high spatial acuity of the fovea. For example, a head-mounted display can consider the gaze direction for visualization of 3D scenes with multiple layers of 2D images. The goal is to improve received image quality and accuracy of focus cues by taking advantage of the loss of spatial acuity in the periphery without the need for tracking the subject's pupil [Lee18]. Another example is a product (called *Foveator*) that tracks the motion of the pupil and limits high-resolution rendering only in the direction needed. The intended application is for improved virtual reality (VR) experience [see www.inivation.com]. These ideas leverage natural design information to relax requirements of a visual system to avoid providing more than necessary as opposed to using the design of natural fovea to inspire newer designs.

3.2.10 Asynchronous event-based retinas [Liu15a]

Conventional camera systems have pixelated, digitized, and framed pictures for post spatial, temporal, and chromatic processing. Natural vision systems send asynchronous action potentials (spikes) when the neuronal voltage potential exceeds a threshold, which happens at any time, instead of on the leading edge of a digital clock cycle. The information is thus gathered asynchronously, and these information spikes only occur when there is something to cause them. Mimicking this biological behavior is the emerging asynchronous event-based systems. Progress has been slow due in part to the unfamiliarity of silicon industry with non-clocked (asynchronous) circuitry. The emulation of cell types is limiting as industry is reluctant to reduce the pixel fill areas to make room for additional functionality [Liu15a]. In mammals the light travels through the retinal neuron layers and then through many layers of photopigment, allowing numerous opportunities for photon capture than a single pass (such as the depletion region of a *pn* junction). The chip real-estate for asynchronous (analog) processing in the retina does not conflict with the photon-capturing photoreceptor as the retinal is transparent to the incoming photonic information.

One asynchronous silicon retina design includes an attempt to mimic the magno- and parvo-cellular pathways (MP and PP) of the optic nerve [Zag04]. The sustained nature of the PP and the transient nature of the MP is pursued which results in both ON and OFF ganglion cells for both MP and PP, which is what is observed in natural vision systems. The benefit is natural contrast adaptation in addition to adaptive spatio-temporal filtering. The resulting localized automatic gain control provides a wide dynamic range and makes available two separate spatial-temporal bandpass filtered representation of the image. One of the challenges of using this vision system is the large non-uniformity between pixel responses [Liu15a]. Gross non-uniformity between receptors and neurons is common in natural sensory systems as the adaptive (plastic) nature of neurons compensates for such non-uniformities.

3.2.11 Emulating retina cells using photonic networks of spiking lasers [Rob20]

Silicon retinas and cochleae such as those in [Liu15] use hard silicon to emulate the behavior of biological neuronal networks that are adaptive and exhibit plasticity. Nevertheless, these bio-inspired designs show promise of the applications of such novel sensory systems. In a similar way vertical cavity surface emitting lasers (VCSELs) are used to emulate responses of certain neurons in the retina and are referred to as *VCSEL-neurons*. In biology the photonic energy is converted to a graded (or analog) potential by the biochemistry of the photopigments of the photoreceptor cells. By keeping the information photonic the speeds of computations can exceed 7 orders of magnitude improvement. This dramatic improvement in

information processing performance has wide applications for computationally-intense algorithm frameworks such as artificial intelligence and deep learning.

This effort demonstrates retinal cell emulation using off-the-shelf VCSEL components operating at conventional telecom wavelengths. The VCSEL-neurons were configured to emulate the spiking behavior of ON and OFF bipolar cells as well as retinal ganglion cells. In these silicon and photonic applications, we see biology as an inspiration for novel information processing strategies but then combine those strategies with available technology that does not emulate the way biology works. A similar example of this concept is also seen when the fovea is emulated in the next few applications.

3.2.12 Integrating insect vision polarization with other vision principles [Giak18]

The visual sensory systems of many species are designed to process environment-provided stimulus that have space, time, and color dimensions. Arthropods and some marine species have been shown to have sensitivities to the polarization of light as well. For example, the octopus retina has tightly packed photoreceptor outer segments composed of microvilli that are at right angles to the microvilli of neighboring photoreceptor cells. The microvilli orientation alternates between these right angles and this is believed to give the octopus sensitivity to polarized light [Smith00].

Aluminum nanowire polarization filters are used in this effort [Giak18] to emulate the microvilli of ommatidium, the components that make up the compound eye. Polarization measurements were made to characterize the polarization of several polymers. A previously designed neuromorphic camera system is used with polarization filters to show improvement in visually recognizing a rotating blade if polarization information is used [Giak18].

Chapter 3 Questions

1. Differentiate between passive and active sensors.
2. What is the energy in a photon?
3. How are chemo-reception and photo-reception similar?
4. Describe the three most significant imperfections in biological vision systems and what causes them.
5. Discuss the relationship between connectivity and spatial and temporal acuity
6. What is coarse coding?
7. What are the three information domains in which vision systems extract environmental information?
8. Describe the three major compound eye designs.
9. Give some examples of visual scanning systems in the animal kingdom. What are the advantages and disadvantages of such a system?
10. Why is the retina considered a part of the brain, since the two organs are separated by distance and other components (optic nerve, LGN, etc.)?
11. What are the anatomical similarities between the retina, LGN, and the brain?
12. Explain the serial/planar duality that exists in biological vision systems.
13. Describe the encoding and decoding levels (in orders of magnitude) in the various organs within the primate vision system.
14. Name the five major cell types (layers) in the retina. Which three are connected to the triad synapse?
15. Give the three primary vision information channels in primate vision.
16. DoG or LoG filters are primarily used to model what part of the vision system?
17. What are the commonalities in color vision models concerning luminance and color?
18. What is the photoreceptor mosaic, and how is that like an artistic mosaic?
19. What is the difference between LoG and DoG filters?
20. Discuss degrees of freedom with LoG and DoG filters?
21. Compare and contrast vision system pathways with a conventional wavelet filter bank.
22. How is coarse coding manifested in the vision system?
23. When contemplating a new communication encoding scheme, it is very important to choose an orthogonal basis. But a typical biological set of basis functions are *not* mutually orthogonal. What is the implication?
24. Why are we so interested in biology if natural basis functions are not orthogonal?

Chapter 3 References

[Boyn60] Boynton, R., "Theory of Color Vision", *Journal of the OSA*, Vol. 50, No. 10, pp. 929–944, 1960.

[Brooks18] Brooks, G., "Gaussian-based filters for elementary motion detector delay element: Modeling spatio-temporal pathways with elementary motion detection", *IEEE Research and Applications of Photonics in Defense*, Shalimar, Florida, 2018.

[Chittka96] Chittka, L., "Optimal sets of color receptors and opponent processing for coding of natural objects in insect vision", *Journal of Theoretical Biology*, Vol. 181, pp. 179–196, 1996.

[Daug88] Daugman, J., "Complete discrete 2-D Gabor transforms by neural networks for image analysis and compression", *IEEE Trans. On Acoustic, Speech, and Signal Processing*, Vol. 36, No. 7, 1988.

[Del94] Delbruck, T. and Mead, C., "Analog VLSI adaptive, logarithmic, wide-dynamic-range photoreceptor", *IEEE Int. Symp. Circuits Syst*, pp. 339–342, 1994.

[DeV88] DeValois, R. and DeValois, K., *Spatial Vision*, Oxford University Press, New York, 1988.

[DeV96] DeValois, R. and DeValois, K., "On 'A three-stage color model'", *Vision Research*, Vol. 36, No. 6, pp. 833–836, 1996.

[Dowl87] Dowling, J. E., *The Retina: An Approachable Part of the Brain*, Harvard University Press, Cambridge, Massachusetts, 1987.

[Dup18] Dubeyroux, J., et al., *M²APix: a bio-inspired auto-adaptive visual sensor for robust ground height estimation*, IEEE, ISBN: 978-1-5386-4881-0, 2018.

[Giak18] Giakos, G., et al., "Integration of Bioinspired Vision Principles Towards the Design of Autonomous Guidance, Navigation, and Control Systems", *9th Int. Conf. on Information, Intelligence, Systems and App. (IISA)*, pp. 1–8, DOI: 10.1109/IISA.2018.8633643, 2018.

[Guth91] Guth, L. S., "Model for color vision and light adaptation", *Journal of the Optical Society of America*, Vol. 8, No. 6, 1991.

[Guth96] Guth, L. S., "Comments on 'A multi-stage color model'", *Vision Research*, Vol. 36, No. 6, pp. 831–833, 1996.

[Hass56] Hassenstein, B. and Reichardt, W., "System-theoretische Analyse der Zeit-, Reihenfolgen und Vorzeichenauswertung bei der Bewegungsperzeption des Russelkafers Chlorophanus", *Zeitsschrift fur Naturforschung, Teil* B, Vol. 11, pp. 513–524, 1956.

[Lee18] Lee, S., et al., "Foveated retinal optimization for see-through near-eye multi-layer displays", *IEEE Access*, DOI: 10.1109/ACCESS.2017.2782219, Feb. 14, 2018.

[Liu15a] Chapter 3, "Silicon Retinas", in Liu, S., et al., *Event-based Neuromorphic Systems*, John Wiley & Sons, ISBN: 978-0470018491, 2015.

[Liu15b] Chapter 4, "Silicon Cochleas", in Liu, S., et al., *Event-based Neuromorphic Systems*, John Wiley & Sons, ISBN: 978-0470018491, 2015.

[Lyon89] Lyon, R. and Mead, C., Ch 16 "Electronic Cochlea", in Mead, C, Ed., *Analog VLSI and Neural Systems*, Addison-Wesley, ISBN: 0-201-05992-4, 1989.

[Maf15] Mafrica, S., et al., "A bio-inspired analog silicon retina with Michaelis-Menten auto-adaptive pixels sensitive to small and large changes in light", *Optics Express*, Vol. 23, No. 5, p. 5614, 2015.

[Maha89] Mahawald, M. and Mead, C., Ch 15 "Silicon Retina", in Mead, C, Ed., *Analog VLSI and Neural Systems*, Addison-Wesley, ISBN: 0-201-05992-4, 1989.

[Maha91] Mahawald, M., "Silicon retinal with adaptive photoreceptors", *SPIE Visual Information Processing: From Neurons to Chips*, Vol. 1473, No. 4, Orlando, April 1991.

[Marr82] Marr, D., *Vision*, W. H. Freeman and Company, New York, 1982.

[Mart94] Martinez-Uriegas, E., "Chromatic-achromatic multiplexing in human color vision", in Kelly, D., Ed., *Visual Science and Engineering: Models and Applications*, Marcel Dekker, Inc., 1994.

[Rob20] Robertson, J., et al., "Toward neuromorphic photonic networks of ultrafast spiking neurons", *IEEE Journal of Selected Topics in Quantum Electronics*, Vol. 26, No. 1, 2020.

[Roub12] Roubieu, F. L., Serres, J., Franceschini, N., Ruffier, F., and Viollet, S., "A fully-autonomous hovercraft inspired by bees: Wall following and speed control in straight and tapered corridors", *2012 IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Guangzhou, 2012, pp. 1311–1318, DOI: 10.1109/ROBIO.2012.6491150.

[Smith00] Smith, C. U. M., *Biology of Sensory Systems*, John Wiley, and Sons, ISBN: 0-471-85461-1, 2000.

[Srini02] Srinivasan, M. V., "Visual Flight Control and Navigation in Honeybees: Applications to Robotics," in Ayers, Davis, and Rudolph, Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, Cambridge, Massachusetts, 2002.

[Srini11] Srinivasan, M. V., "Honeybees as a model for the study of visually guided flight, navigation, and biologically inspired robotics," *Physiological Reviews*, Vol. 91, No. 2, pp. 413–460, 2011.

[Strang96] Strang, G. and Nguyen, T., *Wavelets and Filter Banks*, Wellesley-Cambridge Press, ISBN: 0-9614088-7-1, 1996.

[Van17] Vanhoutte, E., et al. "A Quasi-Panoramic Bio-inspired Eye for Flying Parallel to Walls", *IEEE Sensors,* DOI: 10.1109/icsens.2017.8234110, 2017.

[Werb91] Werblin, F. and Teeters, J., "Real-time simulation of the retina allowing visualization of each processing stage", *Proceedings of the SPIE*, Vol. 1472, 1991.

[WikiMM] Wikipedia, "Michaelis–Menten kinetics", Oct. 2020.

[Wu12] Wu, H., et al., "Insect-inspired high-speed motion vision system for robot control"*, Biological Cybernetics*, Vol. 106, pp. 453–463, DOI: 10.1007/s00422-012-0509-3, 2012.

[Zag04] Zaghloul, K. and Boahen, K., "Optic nerve signals in a neuromorphic chip I: Outer and inner retina models", *IEEE Trans. Biomed. Eng.*, Vol. 51, No. 4, pp. 657–666, 2004.

Chapter 4: MECHANO-SENSORY SYSTEMS

4.1 Natural Mechano-sensory Systems

The primary mechano-sensory systems provide the sense of touch and hearing. Neurons are stimulated by contact with objects in the environment or by contact with fluid compression waves caused by movements in the atmosphere or underwater. The primate auditory sense is caused by air vibrations against the eardrum, which causes bone vibrations in the inner ear, which causes deformations of the basilar membrane that resemble the shape of the frequency spectrum of the incoming sound energy.

4.1.1 Mechano-sensory capability in simple life-forms

The most basic sense is the mechano-sensory *tactile* sense, which is the response to mechanical distortion. The history of the tactile sense goes back to ancient *prokaryocytes*, which are cellular organisms with no distinct nuclei, such as bacteria or blue-green algae. For these fundamental life forms, the tactile sense is required for continually detecting the continuity of the cell boundaries. The organism can then 1) counteract swelling due to *osmotic* forces (fluid entering the cell to balance ionic concentrations) and 2) prepare for cell division when the tactile sense detects swelling for that purpose. [Smith08].

4.1.2 Mechano-sensory internal capability within higher life forms

The human hypothalamus located under the brain serves as an interface between the nervous system and the *endocrine* (interior secretion) system. The fluid secretions controlled by the hypothalamus are a primary influence for heart rate and other biological rhythm control, body temperature control, hunger and thirst control, digestion rate, and other related functions involving secretions. It is believed to be the center for "mind-over-body" control as well as for feelings such as rage and aggression [Tort84]. Within the hypothalamus is a complex neuronal design based on stretch-sensitive mechanoreceptors that sample the conditions of blood cell membranes in an analogous way that they serve the single-celled organisms. The difference is that the prokarycyte stretch-sensitive mechanoreceptors built into the organism, while the hypothalamic mechanoreceptors sample the blood cells from external to the cell [Smith08].

Mechano-sensors are built around stretch-sensitive channels that allow immediate detection and rapid response. Photo-sensory and chemo-sensory reception involves a complex biochemistry to translate the presence of a photon or a chemical tastant or odorant into an ionic charge presence within the receptor. The ionic charge increase is then translated into nerve impulses to eventually be processed by the higher brain functions. The mechanoreceptors, on the other hand, respond immediately to mechanical distortion.

4.1.3 The sense of touch

Mechanoreceptors are fundamental to the detection of tension and the sense of touch. They are also basic components to detecting vibrations, accelerations, sound, body movement and body positions. They play an important role in *kinesthesia*, which is sensing the relative positions of different body parts. It is believed that all these senses are ultimately derived from stretch-sensitive channels. However, the human understanding of the molecular structure and nature of most mechanosensory channels is still in its infancy.

4.1.4 Mechano-sensory sensilla

A discriminating characteristic of arthropods is their external skeleton, which limits (fortunately!) their overall size. Sensory organs such as the retina cannot develop in the hard exoskeletons. However, their kinesthetic sense is well developed due to sensory endings in muscles and joints of appendages. The most common insect sensory unit is the mechanosensory *sensilla*, each of which includes one or more neurosensory cells within a *cuticular* (external shell) housing. The *cuticle* is the external surface of an arthropod. Mechanosensitive sensilla may respond to cuticular joint movements or be positioned to detect movements within the cavity. The three primary mechano-sensory sensilla in arthropods include:

Hair sensilla

Neurosensory cells have dendritic inputs from within a hair protruding from the cuticle and axonal outputs from the cell bodies located at the root of the hair embedded in the epidermis underneath the cuticle. Deflection in one direction causes depolarization (an increase from the nominal –70 mV resting potential) while deflection in the other direction causes hyperpolarization (a decrease from –70 mV.) Minimum response thresholds are known for distortions down to 3–5 nm with response times down to 100 us (0.1 ms). These figures imply the use of opening and shutting ion gates; the biophysics for mammalian hair cells is similar.

Campaniform sensilla

The hair has been reduced to a dome on the cuticle. The dendritic inputs are just beneath the external surface so that the neurosensory cell senses a slight surface deformation. Responses have been shown with deformations as small as 0.1 nm. Directional selectivity is achieved with elliptically shaped domes, where deformation along the short axis is more sensitive to deformation along the long axis.

Chordotonal organs – Mechanosensory sensilla developed within the body cavity. These are characterized by a cap or other cell that stimulates numerous dendritic inputs to the neurosensory cell. Chordotonal organs are one of the proprioceptor types, located in almost every exoskeletal joint and between body segments. Many are sensitive to vibrations; for example, one type in the cockroach is sensitive to vibrations between 1 kHz to 5 kHz and amplitudes from 1 nm to 100 nm (0.1 micron). These sensing capabilities are important for detection of danger and for social communications.

Other non-mechanosensory sensilla include gustatory (taste), olfactory (smell), hygroscopic (humidity-sensing), and thermal (temperature-sensing) sensilla.

Separating insect mechanoreceptors into vibration detectors and acoustic detectors is difficult since many times the same receptors are used to detect vibrations in air, water, and ground. Certain water insects (pond skater, *Gerris*, and water-boatman, *Notonecta*) detect wave amplitudes around 0.5 microns in a frequency range 20–200Hz and time delay range 1 to 4 ms.

Hairs and tympanic membranes for auditory sensing

Two basic types of sound detectors have been developed in insects: Hairs and tympanic organs. Hairs only respond to lateral distortions of air when the insect is very near the sound source, such as the wing beat frequency of a predator insect or a prospective mate. They are accompanied in detecting vibrations by Johnston's organ, which consists of largely packed sensilla. The Johnston's organs also detect flight speed in bees and gravity in the water beetle.

Tympanic organs (ears) respond to pressure waves and are thus able to respond to sound sources much farther away. Tympanic organs are used for communications, attack, and defense. The basic parts include a tympanic membrane, air cavity, and a group of chordotonal organs that provide the neuronal signaling from the acoustic stimulus. Across the species the tympanic organs have developed on many different parts of the insect body.

Evasive maneuvers of the lacewing moth

An interesting use of the tympanic organ is found in the green lacewing (*Chrysopa carnea*). Military pilots being pursued by enemy aircraft may mimic the lacewing when being pursued by a hungry bat. As the bat detects its prey and closes in, its active sonar pulses increase in frequency. When the search pulses are detected, the lacewing folds its wings into a nose-dive out of the sky before the bat's sonar can lock on. Noctuid moths have two neurons for each tympanic organ. One signals a bats detection sonar pulse while the other starts responding with the higher frequency tracking pulses. With the first signal the moth will retreat in the opposite direction; with the second signal it will try desperate avoidance maneuvers, such as zig-zags, loops, spirals, dives, or falling into cluttering foliage. (Surrounding nearby vegetation "clutters" the returning sonar pulses echoing off a target moth; similarly, vegetation also "clutters" the returning echo radar pulses echoing off a military target.) Some moths will emit sounds during the last fraction of a second; it is not sure if the moth is warning others or trying to 'jam' the bats echolocation analysis mechanism [Smith08].

Equilibrium and halteres

Hair cells in different orientations lead to gravitational force detection from different orientations, which lead to balance and equilibrium. Fluid-filled tubes in the vertebrates called the *semicircular canals* are oriented orthogonal to each other. Two fluids, called *endolymph* and *perilymph* are very different with respect to ionic concentration levels. K^+ ions flow through the stereocilia, which project well into the K^+-rich endolymph. The resulting design is a complex system of orientation signals that are processed to achieve balance and equilibrium.

The membranous labyrinth has developed from early lamprey (eel-like fish). It includes the semicircular canals and fluid-filled chambers called the *utriculus* and *sacculus.* It also includes pre-cochlear organs and cochlea (auditory part of hearing system) in the higher species.

Many insects have two pairs of wings to help control their flight, but the *dipteran* (two-winged) insects have developed *halteres* to replace the hind wings. These organs are attached to the thorax just under each wing and have dumbbell shaped endings causing responses to changes in momentum. Dipteran insects typically have short, stubby bodies that make it particularly remarkable that they can control their flight. The halteres provide inertial navigation information that is combined with optic flow input through the vision system. The head is kept stabilized by its own visual input, while the halteres provide inertial information used to stabilize flight. The halteres can be thought of as vibrating gyroscopes that serve as angular rate sensors [North01]. It can be shown that a system of two masses suspended on a stiff beam at 45° has the capability to provide sufficient information for stabilized flight control. How the neurons are connected and how the information is processed to accomplish stabilized flight control, however, will remain a mystery for a long time to come [North01].

The halteres have numerous campaniform sensilla nerve endings attached at the end as well as numerous chordotonal organs embedded within. These signals can detect slight motion in each of the three degrees of freedom: *pitch, roll,* and *yaw. Pitch* is rotation about a horizontal axis orthogonal to the main horizontal axis, *roll* is rotation about the main horizontal axis, and *yaw* is rotation about the vertical axis. To illustrate each of these three, consider the effects of rotational motion when looking ahead from the bow of a ship: Pitch causes up and down motion, roll causes the left side to go up when the right goes down (and vice versa), and yaw causes the ships heading to oscillate to the left and right. Halteres can oscillate through about 180° at frequencies between 100Hz and 500Hz [Smith08].

4.1.5 Mammalian tactile receptors

In mammalian skin tactile receptors can be classified into fast adapting, which respond only during initial skin deformation, and slow adapting, which continue to respond if the deformation is present. Fast adapters include:

- *Pacinian corpuscles*, which are in the deeper layers of glabrous (non-hairy, like the palm) skin and respond to vibrations in the range of 70–1000 Hz
- *Meissner's corpuscles*, which are also in the deeper layers of glabrous skin and respond to vibrations in the range of 10–200 Hz
- *Krause's end bulbs*, like Meissner's corpuscles but found in non-primates, responding to vibrations in the range of 10–100 Hz
- *Hair follicle receptors*, which are located just below the sebaceous (sweat) glands; numerous nerve endings give hair follicles a wide range of hair-movement sensitivities and response times.

The slow adapting tactile receptors in mammalian skin include:

- *Merkel cells*, which respond to sudden displacements, such as stroking
- *Ruffini endings*, which respond to steady displacement of skin
- *C-Mechanoreceptors*, located just beneath the skin, in the epidermis/dermis interface, have unmyelinated (unprotected) nerve fibers extending into the *epidermis* (the most external layer of skin). These nerves respond with slowly-adapting discharge to steady indentations of the skin. They also respond to temperature extremes and to tissue damage, interpreted as pain.

Basic hair cells are similar in structure among all vertebrates. Peak sensitivities in the human ear reach movements of only a tenth of a nanometer, which is one angstrom. Hair cell sensitivity "is limited by the random roar of Brownian motion" [Smith08]. Hair cell ending are composed of bundles of fine hair-like bundles called *stereocilia* and a single, tall cilium with a bulbous tip called a *kinocilium*. The receptor potential depolarizes (rises from–70 mV) for motion in one direction and hyperpolarizes (decreases below –70 mV) for motion in the other direction. (Biologists refer to the normal neuronal resting potential of –70 mV as the natural voltage "polarization" state).

4.1.6 Human auditory system

The founder of Ohm's Law, G. S. Ohm, once suggested that the human auditory systems does a Fourier analysis of received sound signals, breaking the signals into separate components with separate frequencies and phases [Kand81]. Although this has proven to be true, the auditory system does more than a simple Fourier analysis. The input is fluid pressure waves (sound in air) from the environment striking the eardrum and the ear transforms the pressure waves to neuronal signals processed by the auditory cortex in the brain.

Figure 4.1.6–1 shows a sketch of the key components of the human auditory system. Sound enters the outer ear, and the vibrations are transferred to the middle ear and then the inner ear. The outer ear is composed of the external cartilage, called *pinna*, the ear canal, and the tympanic membrane, or *ear drum*. The middle ear is composed of three bones in an air-filled chamber; the inner ear, or *membranous labyrinth*, contains the *semicircular canals*, fluid-filled chambers called the *utriculus* and *sacculus*, which are near the semicircular canals (but not labeled in Figure 4.2.1–1) and the cochlea.

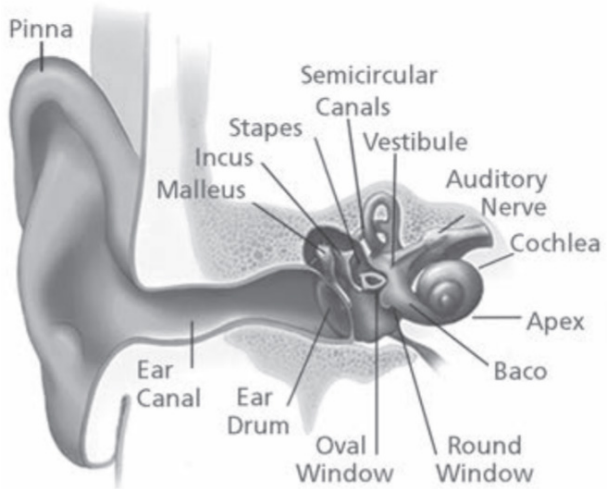


Figure 4.1.6–1 Human Auditory System.

Credit: NIH Medical Arts, Picture from https://www.nidcd.nih.gov/sites/default/files/Documents/health/hearing/AgeRelatedHearingLoss.pdf

The outer ear is designed to collect sound waves and direct them into the ear canal to the eardrum. The *middle ear ossicles* are the *malleus*, or "hammer (mallet)", the *incus*, or "anvil", and the *stapes*, or "stirrups". The names come from their shapes being similar to familiar objects. The ossicles serve to provide an acoustic impedance matcher between the air waves striking the eardrum and the fluid waves emanating from the oval window in the cochlea. Without the impedance matching most of the air-wave energy would reflect off the surface of the cochlear fluid. Another purpose of the ossicles is to amplify the energy density due to the variation in acoustic surface area: The eardrum surface area is about 25 times larger than that of the oval window.

Time delays and sound localization

For humans, the typical time delay for a sound wave to reach each eardrum is between 350 to 650 microseconds [Mead89], depending on the binaural separation distance. A source directly in front of the listener will reach each ear simultaneously with no time delay, while a source at right angles will reach each ear with this maximum time delay. The difference in wave-front arrival time is therefore one of the horizontal localization cues for the sound source, as will be shown later for the barn owl.

Another horizontal localization cue for humans is the result of high frequency attenuation caused by sound traveling around the head. This is referred to as the *acoustic head shadow*. A sound source from directly ahead will have the same attenuation effect in both channels, while a source coming from an angle will result in more high frequency attenuation at the *contra-lateral* (opposite-sided) ear. The sound impulse response from a source between center and right angles shows both a delay and a broadening on the contra-lateral ear with respect to the *ipsi-lateral* (same-sided) ear.

Elevation information is encoded in the deconstructive interference pattern of incoming sound wavefronts as they pass through the outer ear along two separate paths: The first path is directly into the ear canal, and the second is a reflected path off the pinna (see Figure 4.2.1–1) and again off the *tragus* before entering the ear canal. The tragus is an external lobe like the pinna but much smaller (and not seen in Figure 4.2.1–1); the tragus is easily felt when the finger is at the opening of the ear canal. The delay time in the indirect pinna-tragus path is a monotonic function of the elevation of the sound source. Since the destructive interference pattern is a function of the delay time, this pattern serves as a cue for the elevation of the sound source with respect to the individual.

Static and dynamic equilibrium

The three semicircular canals are mutually orthogonal to make available signals from each degree-of-freedom. Two chambers connect to the canals, called the *utricle* and *saccule*. Static equilibrium is sensed in regions of the chambers, while dynamic equilibrium is sensed at the *cristae* located at the ends of the semicircular canals.

The *macula* in the utricle and saccule (inner ear chambers) serve to provide static equilibrium signals. Hair cells and supporting cells in the macula have stereocilia and kinocilium extending into a gelatinous layer supporting *otoliths* (*oto* ear, *lithos* stone). The otoliths are made of dense calcium carbonate crystals that move across gelatinous layer in response to differential gravitational forces caused by changes in head position. The movement stimulates the hair cells that provide static equilibrium signals to the vestibular cochlear nerve. (The vestibular branch contains signals from the semicircular canals and the utricle and saccule chambers, while the cochlear branch contains signals from the cochlea).

The *cristae* located in the ends of each semicircular canal serve to provide dynamic equilibrium signals. Head movements cause endolymph to flow over gelatinous material called the *cupula*. When each cupula moves it stimulates hair cells comprising the *ampullar* nerve at the end of each of the semicircular canals. These signals eventually cause muscular contractions that help to maintain body balance in new positions.

Time-to-frequency transformation in the cochlea

Sound vibrations from the external environment strike the eardrum, causing a chain reaction through the middle-ear ossicles that transforms the air vibrations into fluid vibrations in the basilar membrane of the cochlea. As shown in Figure 4.1.6–2, if the basilar membrane (inside the cochlea) were uncoiled and straightened out, it would measure about 33 mm long, and 0.1 mm (100 microns) wide at the round window end and 0.5 mm (500 microns) wide at the other end [Smith08]:

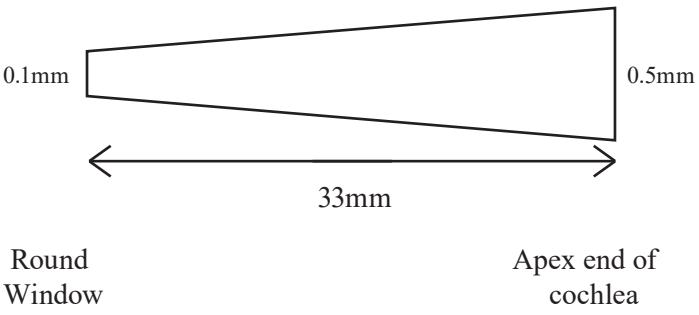


Figure 4.1.6–2. Typical Basilar Membrane Dimensions.

The vertical and horizontal dimensions are scaled differently.

The basilar membrane is stiffer at the round window end and looser at the apex. This causes the wave propagation velocity to slow down as it travels down the basilar membrane. Depending on the initial frequency of the wave, this variable velocity behavior will cause a maximum resonant distortion along the path from the round window to the apex. The basilar membrane is quite complicated and includes sensitive inner and outer hair cell neurons that will respond to deformations of the basilar membrane at the location of each neuron. The hair cell neurons are located along the entire pathway so that the frequency content of the sound can be determined from the spatial location of the neurons that are firing. Thus, the basilar membrane performs a mechanical Fourier Transform on the incoming sound energy and the spatially-distributed neurons sample that signal spectrum.

It was mentioned (Chapter 2) that sensory receptors adjacent to each other in the peripheral sensory system (such as the auditory system) will eventually fire neurons adjacent to each other in the auditory cortex. The relevant signal characteristic of adjacent neurons in the auditory sensor, namely the hair cell neurons adjacent to each other in the basilar membrane, correspond to adjacent frequency components in the input sound. This tonotopic map of the neurons of the basilar membrane is reconstructed in the auditory cortex as well. So, frequency cues are provided by which neurons are firing.

Data sampling rates and coarse coding

The rate of neuronal firing in the cochlea encodes the mechanical distortion of the basilar membrane, which is a direct consequence of the sound energy level of the source. This design is quite remarkable when considering the firing rate of neurons being a maximum of around 1 to 2 ms. Nyquist sampling criteria states that a 1ms sampling (1 kHz) of a signal can only encode information up to 500 Hz, yet human hearing can discern frequencies well above 10 kHz. Each neuron can only fire at a rate much less than the Nyquist criterion, but there are many neurons firing simultaneously, so the aggregate sampling rate is much more than that required to sample a signal whose bandwidth is that of the typical human hearing range (up to 20 kHz).

The firing rate of neurons in the cochlea (basilar membrane) encodes sound intensity information, and not the sound frequency content. The frequency is coarsely-coded: Each neuron has roughly a gaussian frequency response, responding to around 10% to 20% of its peak frequency response. An adjacent neuron would have a slightly different peak frequency response. If both neurons are firing at the same rate, then the frequency would be the value in between their responses. If one is slightly higher than the other, then the frequency component would be closer to its peak response. With only two broadly overlapping gaussian-like frequency responses, a very specific frequency could be extracted with precision far beyond what either neuron could provide.

This is yet another example of coarse coding. In the vision system we observe 4 photoreceptor types whose spectral response curves broadly overlap, yet due to the complex post-processing of highly interconnected neuronal tissue, millions of combinations of color, tone, and shade can typically be discerned. Similarly, the auditory mechanoreceptors are sensitive to frequencies in a 10–20% band around a peak, yet we can discern more specific frequencies at a much higher resolution.

Figure 4.1.6–3 shows a Matlab-generated plot of three Gaussian curves centered at 1.0 KHz, 1.1 KHz, and 1.2 KHz. For a monotone input somewhere between 0.9 KHz and 1.3 KHz would stimulate all three neurons. Keep in mind that the intensity of each neuron response concerns the intensity of the sound, so a moderate response from one neuron could be a weak signal at its peak frequency response, or a stronger signal at a slightly different frequency. For the neuron whose peak response is 1.0KHz, the response would be about the same for a signal at 1.0 KHz, or a 850 Hz signal at twice the strength (where the normalized response is about 0.5), or a 800 Hz signal at 4 times the strength (where the response is about 0.25). A single neuron cannot use its response for very accurate frequency detection.

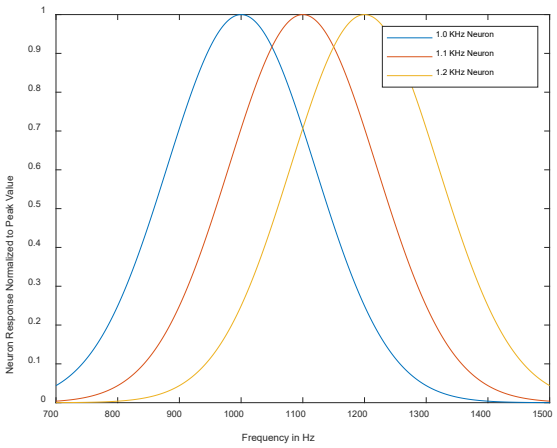


Figure 4.1.6–3. Gaussian curves representing notional auditory neuron responses.

Each neuron will respond to a given input frequency between 900 and 1300 Hz. The specific combination of the responses gives a specific frequency. A single neuron can only give a range of possible frequencies.

It is therefore the *relative* responses of adjacent neurons (that are spatially located along the basilar membrane) that provides the frequency queues. The following example and exercise are intended to simply illustrate the improved frequency resolution that is obtained by comparing the responses of adjacent auditory neurons.

Example 4.1.6-1

Assume three auditory neurons have gaussian responses around peak frequencies of 2.0 KHz, 2.1 KHz, and 2.2 KHz like those shown in Figure 4.1.6–3. Assume the three gaussian responses have the same variance but these three different peak values. Give an estimate (or a range) for three separate inputs given that the normalized neuron output is measured as

	2.0 KHz Neuron	2.1 KHz Neuron	2.2 KHz Neuron
Input_1	0.2	0.8	0.2
Input_2	0.4	0.9	0.1
Input_3	0.1	0.9	0.4

Solution:

For this problem we are not concerned with the significance of one particular response value, but instead how the response values compare to those of adjacent neurons. Conveniently the 2.1 KHz Neuron give the strongest response to all three inputs, so the tone would be at least close to 2.1 KHz. Notice for Input_1 that the response to both adjacent neurons is the same (0.2). Since all three curves have the same variance and due to symmetry of Gaussian curves the only possible frequency giving this set of responses would be one at exactly 2.1 KHz.

The input frequency Input_2 would be closer to 2.1 KHz than 2.0 KHz or 2.2 KHz, but since the response of the 2.0 KHz Neuron is greater than that of the 2.2 KHz Neuron the input would be closer to 2.0 KHz than 2.2 KHz, so something less than 2.1 KHz. If the input frequency were the midpoint 2.05 KHz then we would expect the response values for both the 2.0 KHz Neuron and the 2.1 KHz neuron to be the same, but that is not the case. So, the Input_2 frequency should be greater than 2.05 KHz but less than 2.1 KHz, or in the range of about 2.06 KHz to 2.09 KHz.

The input frequency Input_3 would be closer to 2.1 KHz than 2.0 KHz or 2.2 KHz, but in this case the response of the 2.2 KHz Neuron is greater than that of the 2.0 KHz Neuron, so the input would be closer to 2.2 KHz than 2.0 KHz, so something greater than 2.1 KHz. In this case if the input frequency were the midpoint 2.15 KHz then we would expect the response values for both the 2.1 KHz Neuron and the 2.2 KHz neuron to be the same, but once again that is not the case. So, the Input_2 frequency should be greater than 2.1 KHz but less than 2.15 KHz, or in the range of about 2.11 KHz to 2.14 KHz.

The following give a summary of our estimates for the tonal input frequencies:

	2.0 KHz Neuron	2.1 KHz Neuron	2.2 KHz Neuron	Estimated tonal frequency (KHz)
Input_1	0.2	0.8	0.2	$f \approx 2.1$ KHz
Input_2	0.4	0.9	0.1	$\sim 2.06 \leq f \leq 2.09$
Input_3	0.1	0.9	0.4	$\sim 2.11 \leq f \leq 2.14$

Exercise 4.1.6–1

Assume four auditory neurons have gaussian responses around peak frequencies of 3.0 KHz, 3.1 KHz, 3.2 KHz, and 3.3 KHz like those shown in Figure 4.1.2–5. Assume the four gaussian responses have the same variance but these four different peak values. Give an estimate (or a range) for three separate inputs given that the normalized neuron output is measured as

	3.0 KHz Neuron	3.1 KHz Neuron	3.2 KHz Neuron	3.3 KHz Neuron
Input_1	0.1	0.8	0.8	0.1
Input_2	0.4	0.9	0.8	0.2
Input_3	0.7	0.4	0.2	0.1

Answers:

Input_1: $f \approx 3.15$ KHz,
Input_2: $\sim 3.11 \leq f \leq 3.14$ KHz, and
Input_3: $f \leq 3.0$ KHz■

4.2 Applications inspired by natural mechano-sensory Systems

There are many potential applications for mechano-sensory systems. As can be seen from the example applications that follow, there are numerous natural paradigms to consider for the inspiration of novel design ideas. For example, barn owls, crickets, bats, dolphins, and primate cochlea represent a sample of designs accomplished by attempting to demonstrate or build mechano-sensory systems based on biological inspiration. There are also many useful applications resulting in a divergence from bio-mimicry, such as transforming photonic energy into sound energy and allowing the organism (blind person) the opportunity to learn how to see based on stimulated auditory cues.

4.2.1 Auditory Pathway of the Barn Owl [Lazz90]

The barn owl localizes its prey by using timing delays between the two ears for determining *azimuth* (angle from directly forward) and intensity variations to determine *elevation* (angle from the horizon) with respect to itself. The result is a conformal mapping onto the *inferior colliculus* (IC) of sound events in auditory space. Each sound source is mapped to a specific location in the IC representing azimuth and elevation with respect to itself [Lazz90].

The auditory signals from the cochlea divide into two primary pathways that eventually meet in the IC. The first is the *intensity pathway* and passes through the *nucleus angularis* (NA), encoding elevation information. This is possible in part due to sound absorption variations caused by feather patterns on the face and neck. The second is the *time-coding pathway* and passes through the *nucleus mag-*

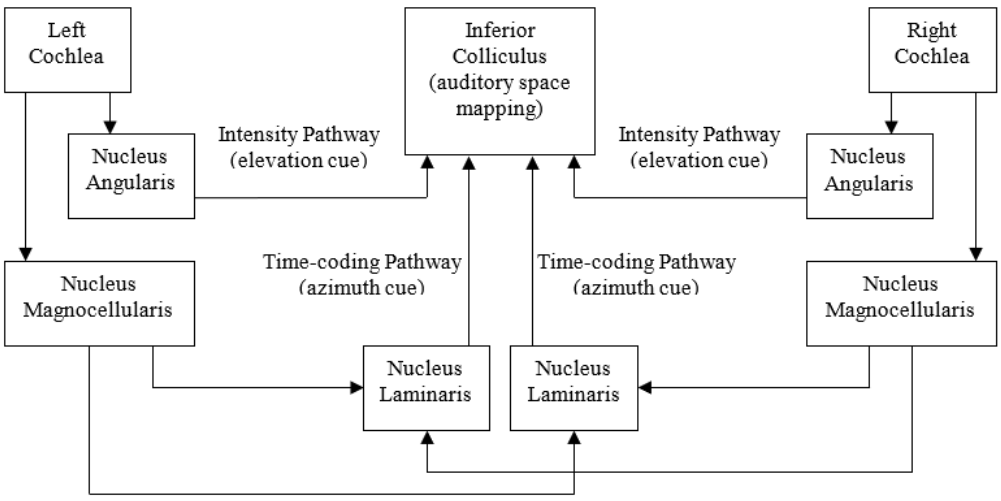


Figure 4.2.1–1. Information Pathways in the Barn Owl.

Sensory pathway for intensity cues is more direct than that for timing cues since the delay from the opposite-side stimulus is used.

nocellularis (NM) onto the *nucleus laminaris* (NL) where it meets the corresponding signals from the time-coding pathway from the opposite side.

Figure 4.2.1–1 represents the two information pathways leading to the IC. The details of the IC are spared to focus more on the pathway structure. Figure 4.2.1–2 shows a notional concept for coincidence detection in the timing circuits of the NL. As drawn, the spatial location of the output signals represents the same spatial direction (azimuth or heading) of the originating sound source.

Assume the total time it takes sound to travel the distance from one ear to the other is divided into 8 time delays, each denoted at Δt, as shown in the model (Figure 4.2.1–2). A stimulus on the immediate left side of the owl (left side of Figure 4.2.1–2) would travel through the bottom row of delays before the right side received the stimulus, therefore resulting in a correlation on the left side. Similarly, a stimulus on the immediate right side of the owl will result in a correlation on the right side of the model. Stimuli in between immediate left or right would result in a correlation somewhere in between these two extremes.

Time-coding Auditory System

The time-coding architecture of the barn owl is implemented in the silicon auditory localization circuit [Lazz90] as shown in Figure 4.2.1–3. Sound enters the system from the left and right ears into respective silicon cochlea described in the previous section. From there 62 equally-spaced taps (representing the basilar membrane neurons in natural cochlea) encode the spectral signature at each side. Each tap feeds a hair-cell circuit that performs half-wave rectification, nonlinear compression, and action potential generation. The action potentials in

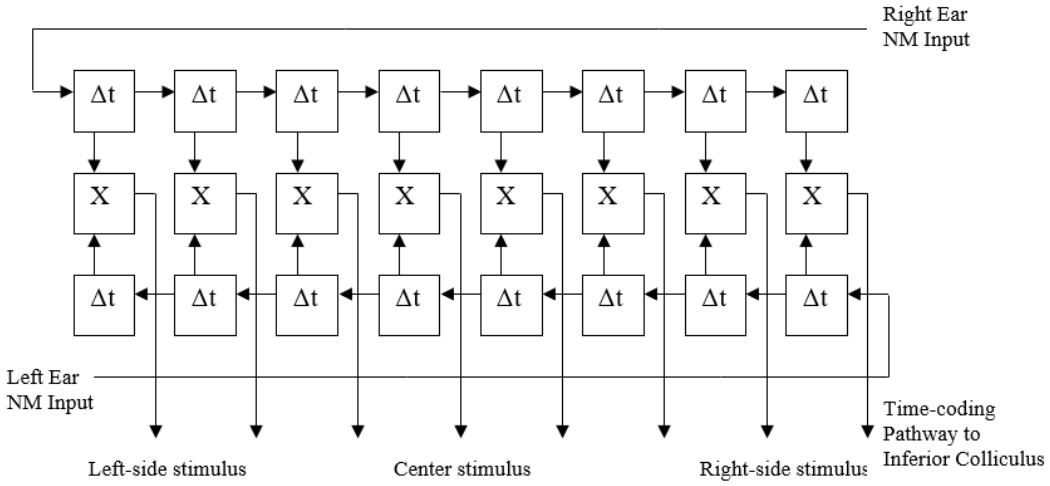


Figure 4.2.1–2. Model of Time-coding in Nucleus Laminaris in the Barn Owl.

Due to delay elements correlations will occur at different stimulus locations.

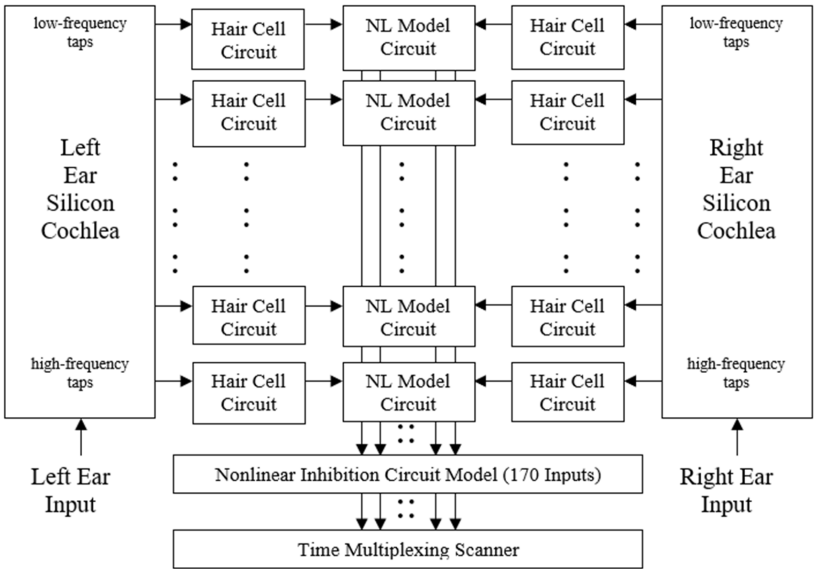


Figure 4.2.1–3. Time-Coding Auditory System Mimics Barn Owl System [LAZZ90].

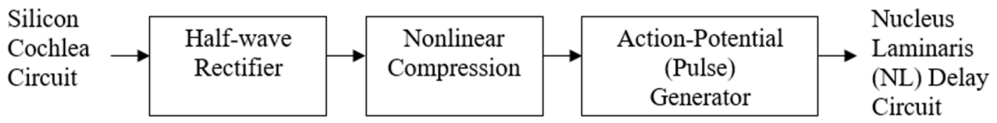


Figure 4.2.1–4. Details of Silicon Barn Owl Hair Cell Circuit of Figure 4.2.1-3 [LAZZ90].

the silicon version are fixed-width fixed-height pulses. As in natural neurons, the frequency of the action potential pulses represents the intensity, and the timing preserves the temporal characteristics of the signal.

The details of the hair-cell circuits are shown in Figure 4.2.1–4. The half-wave rectifier and nonlinear compression simulate the inner hair cells and the action-potential generator simulates the natural *spiral ganglion* cells that take signals from the cochlea in owls, primates, and other species. For the barn owl, these circuits feed the NL-model delay lines like the ones modeled in Figure 4.2.1–2.

4.2.2 Robotic Implementation of Cricket Phonotaxis [Webb01, Webb02]

Cricket Phonotaxis

The male cricket gives a mating call to attract female crickets, and a female can find a specific male using *phonotaxis*, which means movement in response to sound stimulus. In the presence of other noises, the female uses these auditory cues to cover 10 to 20 meters through vegetation and terrain and around obstacles to find the calling male. Phonotaxis is typically seen as a series of start-stop movements with corrective turns.

The "cricket robot" implementing phonotaxis in this example can be modeled as first recognizing the correct song, and then moving toward the source. Each species has a specific sound characterized by a carrier frequency and a temporal repetition structure. A typical pattern is a ten to thirty second syllable of a pure tone (around 4–5 kHz) grouped in distinctive patterns, or chirps. A primary cue serving to discriminate between species is the syllable repetition interval in the song. The correct recognition of this *conspecific* (same species) song is required before migration toward the source.

The cricket does not use time-delay signals between two ears as mammals do nor can it detect phase of the incoming signal. The geometry of the anatomical structure compensates for this inability and gives the cricket the same capability without the complex circuitry. It has an eardrum on each leg connected by an air-filled tracheal tube and two additional openings on the cricket body. Sound reaches each eardrum in two primary paths: one is direct, striking the eardrum on the same side of the cricket as the sound, and the other is indirect, coming from the opposite side of the cricket body. Since these acoustical vibrations are on opposite sides of the eardrum, their effect generally cancels. However, there is a delay due to a longer path-length as well as a delay due to the tracheal tube properties. These delays cause phase differences between the opposing acoustic signals so that the amplitudes do not cancel.

Robotic Implementation

The robotic model of cricket phonotaxis includes a programmable electronic sound source for modeling the cricket call, and a neural network modeling the dynamics of cell membrane potentials. The neural network model is not a generic architecture, but a specific architecture designed to mimic the neuronal structure of the cricket more closely:

> *"The architectures represent neural processes at appropriate levels of detail rather than using standard artificial neural net abstractions. Individual neuron properties and identified connectivity are included, rather than training methods being applied to generic architectures."* [p. 3, Webb01]

The robot is a modification of an existing miniature robot (*Khepera*, "K-team 1994") that is 6 cm in diameter and 4 cm high. It was chosen as it is closer to cricket size than other available robots, although this size is still much more massive than a cricket. A modification for ears added another 6 cm in height. The robot has 2 drive wheels and 2 castors and is programmed in C on a 68332 processor. Due to processor speed limitations, the neuronal model had to be revised (simplified) to run real time. This is a common theme in biomimetic systems: Although conventional processors are 5 or 6 orders of magnitude faster than biological neurons, we still must make sacrifices in computations to achieve any semblance of real-time biomimicry.

Figure 4.2.2–1 shows the simulated neuronal interconnects for the cricket robot. The separation between the microphone ears can be varied but is set at one-quarter of the mimicked species carrier frequency. Another one-quarter period delay is programmed into the inhibitory connection to simulate the delay in the tracheal tube. The inverter (gain of –1) simulates the opposing effects of the direct

and indirect pathways striking the eardrum on opposite sides. In real crickets, the auditory neuron sends signals to the brain, where the connectivity and functionality are still not yet understood. The robotic model includes membrane potentials that result in action potential (spike) signal generation, but the reduction to four simple neurons was done in the robotic implementation in part to keep the simulation operating in real-time.

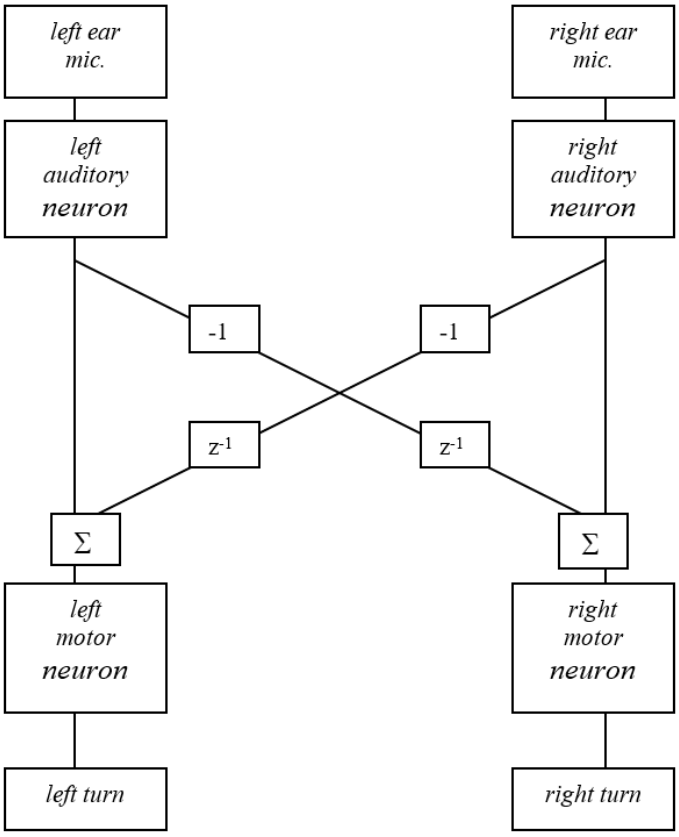


Figure 4.2.2–1. Webb's Neuronal Phonotaxis Model in Cricket Robot. [Webb01]

Microphone ear separation variable, set to a default of ¼ wavelength; indirect inhibitory (gain of –1) pathway with programmable delay (z^{-1}), also set to a default of ¼ wavelength.

Each time a motor neuron in Figure 4.2.2–1 results in an action potential, the robot moves incrementally in that direction. The auditory neurons fire (send action potentials) when the threshold for firing is exceeded. All neurons exhibit leaky integration so that stray noises will not result in action potentials. A constant input stronger than the signal being leaked out must be sustained in order to bring the neuron to firing an action potential. However, the auditory neurons rapidly fire once initiated. This is modeled by returning the membrane potential closer to the threshold (-55 mV typ.) after an action potential instead of returning to the resting potential (-70 mV typ.)

The calling frequency is 4.7 kHz to match a specific species, the *Gryllas bimaculatus*. The robot microphones were placed 18 mm apart, which is a quarter wavelength of the 4.7 kHz calling frequency. An additional one-quarter period delay is also programmed into the circuitry as a 53 us delay. When a signal is received from a right angle to the heading, then the combined delays would add to one-half wavelength, which, when inverted, would combine with the direct signal to give a maximum signal for the motor neuron to turn the robot toward the sound. The opposite motor neuron would receive the direct signal and inverted indirect signal at the same time, thus canceling. When directly in front of the robot, the same signal would be received at both motor neurons so that the left-right turning would cancel, and the robot would continue straight.

Results and discussion [Webb01]

The ¼-wavelength physical ear separation and the ¼-wavelength programmable delay for a 4.7 kHz carrier proved to mimic biological observation. Experimental results showed that the robot migrated toward a 4.7 kHz signal more strongly than a 2.35 kHz signal and would ignore a 9.4 kHz signal. It would also move toward the 4.7 kHz signal when played simultaneously with a 6.7 kHz signal.

By tuning the time constants, the response could be made selective for a bandpass of syllable rates. In one example, the robot responded to changes in signal direction when the syllables were 20 to 30 ms long but would not respond for shorter or longer syllables. The programmability built into this cricket robot will allow further study into the alternate hypotheses of how crickets and other animal species perform phonotaxis. The system will also allow for further study into non-phonotaxis capabilities of such a sensorimotor system.

Although the four-neuron model does not mimic the complexity of the cricket brain, it does demonstrate a minimal configuration for accomplishing basic phonotaxis functions, such as tracking of sound sources, selectivity for specific frequencies, selectivity for syllable rates, tracking behavior without directional input, and tracking behavior in the presence of other sound sources.

4.2.3 Mead/Lyon Silicon Cochlea [Lyon89]

The Mead/Lyon [Lyon89] Silicon Cochlea is a transmission line of second-order amplifier circuits illustrated in Figure 4.2.3–1. First order stages are simple circuits such as differentiators or integrators, whose step responses are typically an exponential response toward a steady-state condition. The second-order stages provide sinusoidal response characteristics to step responses that will provide a

peak response at a resonant frequency. In the initial silicon cochlea circuit, there were 100 second-order circuits with 10 voltage taps evenly spaced along the design.

Each second-order circuit is composed of three op-amps and two capacitors configured as cascaded follower-integrator circuits with a feedback amplifier providing oscillatory responses. The transconductance of the feedback amplifier is controlled by an external bias voltage. For low feedback transconductance, the circuit behaves as a two-stage follower-integrator, which follows the input voltage. As the feedback transconductance is increased, positive feedback causes the second integrator-follower to leap ahead slightly and oscillate to a steady state value. If the transconductance is set too high, the circuit oscillates out of control (goes unstable).

Once appropriately calibrated (tuned), the peak response of each second-order circuit is a function of the input frequency. Since each stage inherently adds a smoothing effect, the individual frequencies of the input voltage signals will have a peak response somewhere along the 100-stage circuit. As in natural cochlea, the spatial distribution of the voltage taps provides a sample of the Fourier representation of the input voltage signal. However, in natural cochlea the mechanical design of the basilar membrane provides physical peak deflections (corresponding to signal frequency components present in the input signal) while this design models the mechanical cochlear structure with a bank of 2nd order electronics filters.

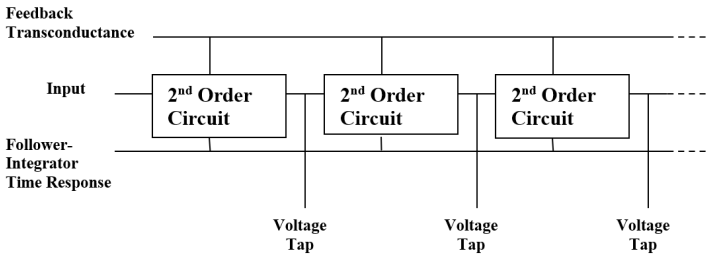


Figure 4.3–1. Silicon Cochlea is delay line of 2nd order circuits. [Lyon89]

4.2.4 MEMS-based electronic cochlea [Andr01]

An example of a Micro-electromechanical system (MEMS) approach to a silicon electronic cochlea is described in [Andr01]. MEMS allows for mechanical distortion due to the incident sound energy to change the distance between two polysilicon plates that are implemented as a capacitor. This design concept includes a MEMS-based acoustic pressure gradient sensor and filter bank that decomposes incident acoustical energy into its wavelet components. The pressure transducer is a conventional MEMS polysilicon diaphragm suspended in air over polysilicon backplate. Inspired by mechanically-coupled acoustic sensory organs of the parasitoid fly, the transducers are connected by a first-layer polysilicon beam, allowing for pressure gradient measurement. As acoustical energy strikes the external plate, the plate is distorted toward the backplate, reducing the air distance separating the two plates. This causes a decrease in the capacitance in response to acoustic pressure. The MEMS silicon cochlea implementation is composed of MEMS filter banks that allow for a real-time wavelet decomposition of the received acoustical energy.

The advantages of the MEMS-based approach over analog VLSI approach is a lower power requirement as the physical energy of the sound waves is doing some of the work of the VLSI transconductance amplifiers. Also, since the MEMS-based approach more closely resembles natural systems there is a more direct correlation with system response to input acoustical energy.

Another application of MEMS technology for biomimetic robots include cantilever microswitches to model antenna behavior and provide water-flow sensors. These MEMS-based sensors are being used to model lobster and scorpion behaviors on underwater robotic vehicles [McGr02].

4.2.5 "See-Hear" design for the blind by retraining auditory system [Mead89]

The "See-Hear" concept is intended to help a blind person "see" by hearing different sounds based on objects visible in a head-mounted camera system [Mead89, Ch 13]. Successful implementation requires transforming visual signals into acoustic signals so that users can create a model of the visual world with their auditory system.

Both vision and auditory systems have receptive fields representing data distributions within the local environment. The vision system maps light emissions and reflections from 3D objects onto the 2D photoreceptor mosaic in the retina, whose conformal mapping onto the brain is called the *retinotopic* map. Similarly, the auditory system takes frequency components of local sound energy and maps a spectrum onto the basilar membrane in the cochlea and subsequently (via cochlear nerve) to a conformal map on the brain called the *tonotopic* map.

Both vision and auditory systems are concerned with detecting transient events. The vision system detects motion by taking time-space derivatives of the light intensity distribution. Transients help to localize events in both space and time, and the brain constructs a 3D model of the world using *motion parallax*, which is the apparent object motion against the background caused by observer

motion. If an observer is focused on a point at infinity and moves slowly, then nearby objects appear to move rapidly against the infinite background, while objects farther away appear to move more slowly. Transient sounds are also easily detected and localized in the auditory system.

The vision and auditory systems differ in how the peripheral information is processed:

> *"In vision, location of a pixel in a 2D array of neurons in the retina corresponds to location of objects in a 2D projection of the visual scene. The location information is preserved through parallel channels by retinotopic mapping. The auditory system, in contrast, has only two input channels; location information is encoded in the temporal patterns of signals in the two cochleae. These temporal patterns provide the cues that the higher auditory centers use to build a 2D representation of the acoustic environment, similar to the visual one, in which the position of a neuron corresponds to the location of the stimulus that it detects."* [Mead89]

The key biological vision concepts exploited in the *See-Hear* chip include [Mead 89]:

- Logarithm of light intensity collected at the photoreceptor; using a logarithmic function expands the available dynamic range as compared to a linear function.
- The spatial orientation of light sources (which includes reflected light) is preserved from the photoreceptor mosaic through the retinotopic map
- Depth cues required for mental reconstruction of 3D space are provided by time-derivative signals of the light intensity profile

The key auditory cues for sound localization include:

- Time delay (350–650 microseconds) between ears, providing horizontal placement cue
- Acoustic high-frequency attenuation, providing further horizontal placement cue
- Direct and indirect pathways in the outer ear causing a destructive interference pattern that is a function of elevation, thus providing a vertical placement cue

As in a natural vision system, the *See-Hear* system accepts photonic energy through a lens and focuses the energy onto a 2D array of pixel. (A *pixel* is simply a *picture element*). Each pixel value represents the light coming from a specific direction in the 3D world. The *See-Hear* chip includes local processing at each pixel location.

Each pixel processor responds to the time-derivative of the logarithm of the incident light intensity. The incoming photons of light enter the depletion region of a bipolar junction phototransistor creating electron-hole pairs in quantities proportional to the light intensity. Two diode-connected MOS transistors connected to the emitter cause a voltage drop in response to the logarithm of the light intensity. A MOS transconductance amplifier with nonlinear feedback provides a time-derivative output signal of the pixel processor. Each pixel processor is capacitor-coupled to adjacent pixels so that each pixel processor act as a delay line.

Time-derivative signals propagate in two directions in the electronic cochlea circuit, which results in a mimicry of the time delays between the left and right ears. As seen in Figure 4.2.5–1 a transient event in the left visual field will result in sound on the left side before sound on the right side, which mimics the behavior of sound events in auditory systems. The time delay circuit also filters higher frequencies, so that longer delays result in more attenuation of higher frequencies. This feature therefore models the binaural head-shadow, which is the attenuation of high frequencies as the sound travels around the head. The combined effect of delayed signals and high-frequency attenuation of the delay channels serves to combine both natural horizontal localization cues into one circuit.

Since each pixel processor circuit in the electronic cochlea contains its own photoreceptor circuit, multiple sound sources are processed as a superposition of the individual sources. To model the elevation inputs from the pinna-tragus pathway differences, the see-hear chip contains an additional delay circuit at each end. The 2D image is focused on a 2D array of pixel processors, and the output of each horizontal row is added to a delayed version of itself to model the mixing of the pathways relevant to the elevation of the objects in the image. In this way, the outputs of each row are all summed together to create only two separate sound signals, one for each ear. If two of the same objects were at different elevations within the image, the different pinna-tragus pathway delays at the end of their respective rows will provide the user with an audible queue as to where (in elevation) the object is located.

The user can ultimately learn how to hear a 3D model of the external environment based on what is visually captured with the camera system.

4.2.6 A biomimetic sonar system [Reese94]

A "Biologic Active Sonar System (BASS)" based on echo processing of bats and dolphins was designed to detect and classify mines in shallow water [Reese94]. Front-end filters and nonlinear functions emulating auditory neuronal models were used to obtain high resolution with low frequency sonars (which is another example of coarse coding in natural systems). The intended product of this research is a system implementation into an autonomous underwater vehicle.

Figure 4.2.6–1 shows the block diagram of the BASS processing stages. The band-pass filters (BPF's) have sharp roll-off characteristics at high frequencies and are broad-band, overlapping other channels significantly (coarse coding). This is inspired by natural peripheral auditory processing and provides good time/frequency definition of the signal as well as increases in-band signal-to-noise ratio (SNR).

As in the vision system, the automatic gain control (AGC) allows for covering a much wider dynamic range, which is based on integrate-to-threshold behavior of auditory signals. This sharpens signal onset time, which translates to sharpening range resolution. The half-wave rectifier and sigmoid function is inherent in mammalian auditory processing and serves to sharpen the onset-time and range resolution.

Peak summing and delay provide in-band coherent addition and inter-band signal alignment. This mimics natural biological phase-locked loops and provides pulse compression. The anticipated benefits of such a wide-band low frequency design is longer detection ranges and better target recognition of partially buried mines.

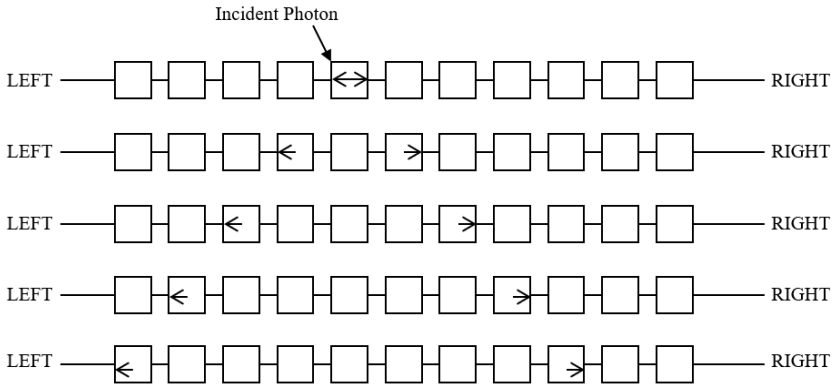


Figure 4.2.5–1. Aural left-right arrival time mimicry in electronic cochlea.

This figure illustrates the part of the electronic which mimics the different arrival times of a sound wavefront due to different pathway lengths to each ear. Each block represents a uniform time-delay circuit and photodetector circuit. Each row in the figure represents the same row at a later unit of time.

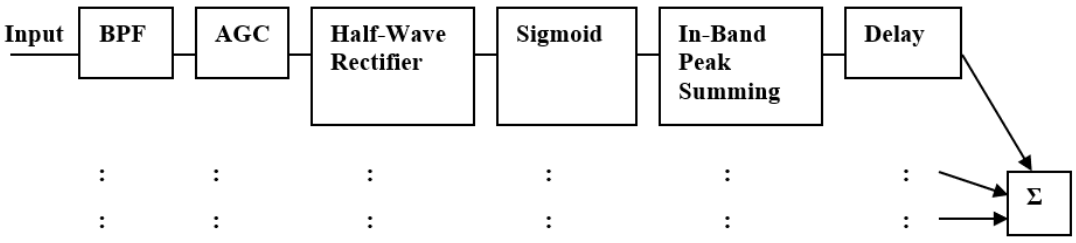


Figure 4.2.6–1. Biologically-inspired sonar system.

Chapter 4 Questions

1. What is the most basic sense?
2. Why is this sense necessary for the most primitive life-forms?
3. How are stretch-sensitive mechanosensory designs fundamentally different from photosensory and chemosensory mechanisms?
4. What is kinesthesia?
5. Lateral inhibition is a form of adaptation. What signal function does it accomplish?
6. How is a military fighter pilot like the green lacewing?
7. How is coarse coding manifested in the human auditory system?
8. What are the three middle-ear ossicles, and what is their function?
9. How are static and dynamic equilibrium changes sensed in the human auditory system?
10. Neurons can fire at most 1kHz, or at a rate of 1ms between action potentials. Nyquist sampling requires two samples per highest-frequency wave period, which means that such a neuronal firing rate can only encode up to 500 Hz. How is it that humans can discern components beyond 20 times that amount (10kHz)?
11. What were some of the significant results from Webb's robotic implementation of cricket phonotaxis?
12. Why is it so amazing that we must "cut corners" in computational processing to get our electronic models to simulate real-time behavior of biological sensory systems?
13. What are the two information pathways in the auditory system of the barn owl?
14. How do first-order systems, such as differentiators and integrators, and second-order systems differ in their step responses?
15. What is the basic idea behind the "See-Hear" system?
16. What advantages does the MEMS-based silicon cochlea have over the analog VLSI-based silicon cochlea?
17. Define these terms:

    phonotaxis –

    pitch –

    roll –

    yaw –

    azimuth –

    elevation –

    halteres –

    dipteran –

    pixel –

    motion parallax –

    MEMS –

Chapter 4 References

[Andr01] Andreou, A. G., et al., "Heterogeneous Integration of Biomimetic Acoustic Microsystems", *IEEE Int. Sym. On Circuits and Systems*, Vol. 3, 2001.

[Ayers02] Ayers, J., Davis, J., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, ISBN: 0-262-01193-X, 2002.

[Kand81] Kandel, E. R. and Schwartz, J. H., *Principles of Neural Sciences*, Elsevier/North-Holland, New York, 1981.

[Lazz90] Lazzaro, J. and Mead, C., "A Silicon Model of Auditory Localization," Chapter 8 in Zornetzer, S., Davis, J., and Lau, C., Eds., *An Introduction to Neural and Electronic Networks*, Academic Press, Inc., ISBN: 0-12-781881-2, 1990.

[Lyon89] Lyon, R. and Mead, C., "Electronic Cochlea", Chapter 16 in Mead, C., Ed., *Analog VLSI and Neural Systems*, Addison-Wesley, ISBN: 0-201-05992-4, 1989.

[McGr02] McGruer, N. E., "Biomimetic Flow and Contact/Bending MEMs Sensors", Chapter 1 in Ayers, J., Davis, J, and Rudolph, A, Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, ISBN: 0-262-01193-X, 2002.

[Mead89] Mead, C., *Analog VLSI and Neural Systems*, Addison-Wesley, ISBN: 0-201-05992-4, 1989.

[North01] Northrop, R. B., *Introduction to Dynamic Modeling of Neuro-Sensory Systems*, CRC Press, ISBN: 0-8493-0814-3, 2001.

[Reese94] Reese, S. and Kenney, J., "Ultra-High Resolution, Biologically-Inspired Sonar", in *Proceedings of the 1994 Symposium on Autonomous Underwater Vehicle Technology*, Ocean Engineering Society and IEEE, Cambridge, Massachusetts, July 1994.

[Smith08] Smith, C. U. M., *Biology of Sensory Systems*, 2nd ed., John Wiley and Sons, ISBN: 978-0-470-51862-5, 2008.

[Tort84] Tortora, G. and Anagnostakos, N., *Principles of Anatomy and Physiology*, 4th ed., Harper & Row Publishers, ISBN: 0-06-046656-1, 1984.

[Webb01] Webb, B., "A Spiking Neuron Controller for Robot Phonotaxix", Chapter 1 in Webb, B., and Consi, T., Eds., *Biorobotics Methods and Applications*, AAAI Press/MIT Press, ISBN 0-262-73141-X, 2001.

[Webb02] Webb, B. and Harrison, R., "Phonotaxis in Crickets and Robots", Chapter 26 in Ayers, J., Davis, J., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, ISBN: 0-262-01193-X, 2002.

Chapter 5: CHEMO-SENSORY SYSTEMS

5.1 Natural Chemo-sensory Systems

Natural chemo-sensory systems provide information from four groups of senses:

General chemical sense: All organisms display this sense. For humans, this sense is mediated by free neurons in the skin.

Olfaction: The sense of smell, generally regarded as a *distance* sense.

Gustation: The sense of taste, generally regarded as a *contact* sense. Separating olfaction and gustation is difficult as the cellular and molecular mechanisms can be the same. We could try to separate the two as either atmospheric or fluid medium, but this breaks down in describing the two senses for underwater life forms.

Solitary chemo-receptor cells (SCCs): Best developed in a few species of fish. The receptors are scattered in the fin surfaces and provide information on the presence of food or predators.

5.1.1 Chemo-sensory capability in simple life-forms

The earliest life-forms on earth were the *prokaryotes*, which are cellular organisms with no nuclei, and the *eukaryotes*, which do have nuclei. It is believed that these organisms had the world to themselves for about two billion years. Much of our understanding of the molecular biology of chemo-sensitivity comes from experiments with contemporary bacteria called *Escherichia coli*, or e. coli.

Moving bacteria are propelled by *flagella*, which are long cilia- or hair-like protrusions that twist or turn in response to chemical stimuli. Some will rotate at around 100 Hz, energized by a transmembrane hydrogen ion concentration gradient. E. coli has 5-10 flagellum that are on either side. When all rotate counter-clockwise, the bacterium moves forward toward a chemical attractant, while when they all rotate clockwise, the motion is a random tumbling motion. With no chemical attractants, the movement is sporadic and random; with attractant present, the motion is the same except there is less tumbling when moving toward the source. The overall motion is a migration toward the source of chemical attractant.

Deep study into certain internal chemosensory system mechanisms will quickly merge into endocrinology (internal secretions) and biochemistry. The olfaction and gustation systems, however, are driven by chemical information *external* to the organism. Our interest is more on these exteroceptor sensory systems than on the interoreceptor sensory driven systems [Smith08].-

5.1.2 Gustation in insects

Chemo-sensory receptors in insects are frequently multi-modal, serving as both a mechano-sensory receptor and a chemo-sensory receptor. The multi-modal sensilla (hairs) protrude from the outer cuticle (shell) with a terminal pore at the tips of the sensilla. Chemicals can enter the pores and travel to the nearby dendritic inputs. Bio-chemical chain reactions result from the combinations of certain chemicals with the nerve endings. These same sensilla would also have another neuron sensitive to the mechanical distortions on the sensilla caused by fluid movement or direct contact or pressure [Smith08].

5.1.3 Gustation in mammals

There are six basic taste qualities [Smith08]:

sweetness

saltiness

sourness

bitterness

umami

water

The first four are in the order of taste receptor cells (TRCs) in the human tongue from the tip and working back. Umami is a Japanese word for the taste of monosodium

glutamate, a crystalline salt used for seasoning foods ($C_5H_8O_4NaN$). Gustatory receptors in mammals are grouped into *taste buds*, which are located on projections called *papillae*. Four types of papillae include

filiform – contains no taste buds; serves to give tongue abrasive character (as in cats)

fungiform – resemble mushrooms; located on front and edges of the tongue; visible red spots sensitive to sweetness and saltiness; buried in the surface epithelium

foliate – located in folds at the rear of the tongue; sensitive to sourness or acidity

circumvallate – sunken in moat or trench; sensitive to sourness or bitterness

From the tip of the tongue to the back, the primary qualities that stimulate the taste buds are in this order:

1) sweetness, 2) saltiness, 3) sourness, and 4) bitterness. Taste Receptor Cells (TRCs) typically have dendrites to multiple taste buds. Similarly, each taste bud may provide input to multiple TRCs. New nerve endings "search out" new synaptic contacts as taste buds are turned over. Thus, there is a complex connection scheme of taste buds to associated TRCs. There is ongoing debate as to whether the brain recognizes different tastes by specific fiber activity or by a pattern of activity across the population of fibers [Smith08].

5.1.4 Olfaction in insects

Insect *hygro-receptors*, which detect humidity, are classed as olfactory (distant receptors) as there is no opening for direct contact to the environment. These sensilla are typically short pegs within a cuticular cavity. Humidity causes sufficient mechanical distortion for receptor signaling, which would explain why they are set within a cuticular cavity: normal contact with the environment will not falsely send a humidity signal.

Hygro-receptors have been detected on the antennas of all insects that have been carefully examined. Although present in all these species, they are typically very sparse among other sensilla. For example, on the cockroach, about one in every 500 sensilla is a hygro-receptor. Hygro-receptor neurons share the same sensilla with other hygro-receptor neurons and with thermo-receptor neurons.

Insect olfactory sensilla are typically multi-porous, allowing extra opportunity for the detection of a *semiochemical*, which is a chemical stimulant, or *pheromone*, with carries a specific meaning, such as a mating opportunity, danger, trail, aggregation, or dispersal. Social insects rely on trails and patches of semiochemicals. The detection of the sex pheromone is the most effective, which makes sense considering the importance of reproduction to survival. A male silkworm moth can detect a single molecule of the female pheromone. A single antenna consists of many branches, each having many sensilla. Each antenna has about 17,000 sensilla that are each 100 microns long and 2 microns in diameter. The large number of sensilla effectively amplify the detection of faint odors in windy conditions.

Olfaction begins with a chemical binding of the attractant molecule to an odorant-binding or pheromone-binding protein. There is increasing evidence that the subsequent biochemistry involving *G-Protein* membrane signaling is the same as found in vertebrate olfactory systems. This suggests a common process that has been developed throughout the animal kingdom [Smith08].

Rheotaxis and Anemotaxis

Insects such as moths use odor-gated *anemotaxis*, which means the insect moves in response to odorants present in the air currents. The moth's flight path is modulated by odor concentrations. One simple strategy for anemotaxis is demonstrated by the male moth moving toward an attractant released by the female moth. When an attractant is detected, male moth flies upwind, and when the odorant plume is lost, it zig-zags across wind, increasing distances. If the male moth detects the attractant again, it simply flies upwind.

Lobsters move there antennas back and forth to detect a source of food underwater. However, lobsters do not use *rheotaxis*, which is basically underwater anemotaxis, since the underwater currents are far too turbulent for that to work.

Their irregular and variable tracks to source and increased speed in middle of track suggests lobsters (and other marine animals) are steered toward plume sources by odor patches, not odor-stimulated up-current movements like the moth. An interesting description is provided by [Consi94]: "Lobsters smell via paired antennules, small antennae positioned medially to the large antennae. Each antennule contains an array of thousands of sensory cells arranged in a tuft of hairs. The antennules can act as discrete time sampling sensors: under conditions of low flow they periodically 'flick', ejecting a parcel of water and allowing a new packet to enter the tuft of sensory hairs for a new measurement."

5.1.5 Olfaction in mammals and other vertebrates

Fish make incredible use of the sense of olfaction. Sharks and dogfish can detect blood and other body fluids from long distances. Salmon can use their olfactory sense to return to their spawning ground by tracing faint chemicals unique to their place of birth.

Receptor field mapping is obvious in the visual, auditory, somatosensory, and (to a lesser extent) the gustatory systems. Olfactory systems do not exhibit a receptive-field mapping corresponding to spatial location of the external environment. It does appear that there are three or four *expression zones*, where each zone represents each of the various types of molecular stimulants.

Individual olfactory receptor cells (ORC's) are tightly embedded between supporting olfactory epithelium cells, with up to 20 cilia that detect stimulants and transmitting action potentials to the next layer of cells, the mitral cells. Photoreceptors in the vision system are also embedded between epithelium cells, but photoreceptors transform photonic flux into graded (analog) signals for processing by the next layers in the retina instead of action potentials. There is a convergence of about 1000 ORC's to one mitral cell, and about 25 mitral cells to form one glomerulus. All 25000 or so ORC's (in the rabbit) that converge to a glomerulus are specialized to detect one (or similar) odorant molecule, so that each glomerulus responds to one specific odor type [Smith 00].

5.1.6 Similarities in vision and olfactory systems: the retina and the olfactory bulb

The table at right is a summary of some similarities between the preprocessing stages in the vertebrate vision and olfactory systems. In both the retinal and the olfactory bulb there are two layers of cells connected orthogonal to the direction of information flow that mediate or inhibit the forward flow. This mediation serves to accent the locations of stimuli within the receptor layer and minimize the signal energy propagated to the deeper neuronal processing layers in the brain.

VISION	OLFACTION	FUNCTION
Retina	Olfactory Bulb	Preprocess Information
Photoreceptors (Rods and Cones) (graded output)	Olfactory Receptor (ORCs) (spiked output)	Receive stimulus
Horizontal	Periglomerular	Mediate (Inhabit) nearby response
Bipolar	Mitral/Tufted	Pass on mediated signal
Amacrine	Granule	Further mediation (?)
Ganglion	Mitral/Tufted	Pass on mediated signal
100:1 (1:2 fovea)(ORC Glom) (400:1 periphery)	1000:1 or 25000:1 (ORC to OT)	Receptor signal compression

5.1.7 Coarse-coding in vision and chemosensory systems

There are a relatively few specialized ORC types from which we can discern many different smells. Specific odors cause specific patterns of responses to ORC types, so odors are analyzed by spatial maps in the central nervous system like the way other distant senses are mapped [Smith08]. A 'model nose' [Persaud82] is discussed later in this chapter where the authors searched for unique patterns of many odorants (over 20) using the responses of only three commercially-available sensors. This demonstrates coarse coding, previously defined as the transformation of raw data using a small number of broadly overlapping filters. The power of coarse coding is that detailed resolution can be achieved with relatively few broadly overlapping sensor responses. A handful of broadly-overlapping sensors can provide raw data for identifying thousands of different categories (smells, tastes, etc.).

Gustation (taste) sensory systems are like olfactory ones in that there are relatively few types of receptor cells whose responses overlap significantly across numerous input types (tastes). A lot of work has gone into understanding psychophysics, or documenting behavioral responses to organism inputs, as well as microbiology for understanding neuronal and other cellular responses to environmental inputs. A significant gap in knowledge exists for explaining how the individual cellular responses are combined and biologically processed to give the overall behavior.

As previously mentioned, in vision systems coarse coding exists in time, space, color, spatial frequency and temporal frequency domains, and here we find in olfactory and gustatory domains as well. With relatively simple (or low-order) filters or sensory responses biology offers a high degree of acuity in these sensory information domains. We saw that in vision systems there are essentially only four broadly-overlapping chromatic detector types, three basic temporal channels, and three basic spatial channels. Neurons receiving broadly-overlapping photoreceptor responses provide higher brain processing areas the ability to discern details in color, time, and spatial domains. Similarly, broadly-overlapping olfactory and gustatory receptor responses provide higher brain processing areas the ability to discern many distinct odors and tastes.

5.2 Applications inspired by natural Chemo-sensory Systems

There are many potential applications for chemo-sensory systems. Some of these include

- Health industry to understand our own biology
- Commercial food industry to understand sense of taste
- Commercial perfume industry to understand sense of smell
- Commercial pesticide/herbicide industry to understand insect gustation
- Governments to monitor air and water and find sources of pollution
- Military to trace chemical trails to hidden explosives, etc.

As with photo-sensory and mechano-sensory system applications, the rest of this chapter represents a sample of contributions of scientists and researchers attempting to demonstrate or build chemo-sensory systems based on biological inspiration.

5.2.1 A model nose demonstrating discrimination capability [Persaud82]

This work demonstrates the coarse-coding capability of an olfactory system. The ORC types have different response levels for the basic odor components, and specific odors are believed to be perceived as a combination of the ORC type responses.

A model for an artificial olfactory system simulating biological ones was pursued with a focus on selecting odorant detectors that respond to a wide variety of chemical types and combining the responses so that different odorants can be identified in parallel. A ratio of sensor responses was used to discriminate between different stimulating odorants.

Gas sensors made by using n-type semiconductors is convenient, as the dopant and intensity of the doping could be adjusted to achieve a desired biomimetic ORC type. This would be advantageous as semiconductor technology is well suited to make a wide variety of n-type semiconductors with very uniform responses, each representing an ORC type. A set of such commercially available semiconductor gas sensors were used and gave specific response patterns for specific stimulants, but the response times were not as fast as natural olfactory systems.

The model nose was completed by using three commercially-available sensors from Figaro (www.figarosensor.com), including one intended as a general purpose combustible gas sensor, one more sensitive to alcohols, and one more sensitive to carbon monoxide. The results showed that the responses to over 20 different odorants were consistent and unique. The researchers point out that as in biological systems, such an artificial olfactory system would have to be trained to recognize specific patterns as specific odors.

5.2.2 Integrating a sniff pump in an artificial olfactory sensor [White02]

The Tuft Medical School Nose (TMSN) was designed to improve sensitivity and discrimination ability with respect to previous artificial nose efforts. A fan and valving system were arranged so that odorant molecules were drawn over the olfactory sensor array in short bursts, mimicking inhalation patterns.

A deviation from biology includes the use of polymer and dye mixtures in LEDs whose fluorescence changes based on the present odorants. So electric energy is used to illuminate LED whose spectra change with input odorants, and then the photonic energy is converted to analog electronic for further processing. Presumably, this is done to help meet the desired sensitivity for a specific application, the one here being land mine detection. This device included 32 sensors whose responses were broad across the various odorants, which included TNT, DNT, and other such compounds. This course-coding of the input resembles natural olfactory sensors.

This project (funded by ONR) illustrates the different uses of biology for inspired design. One purpose is to emulate biology to better understand how biology does what it does, so it is very important to make every effort to not deviate from biology. Another purpose involves a separate

problem that needs to be solved (detecting land mines) where biology can give some incredible insights into novel designs, but to meet the objectives other technology may be integrated int the design that moves it away from true emulation of biology.

5.2.3 Integrating spike-based processing into artificial olfactory sensor [Liu18]

This effort contributes the integration of spike-based signal processing which is a known characteristic in natural olfactory sensors. The first sensing stage is an array of *virtual olfactory receptor neurons* (VORNs) that convert the odorant response into a spatio-temporal pattern of spikes. As in biological ORs the array is composed of groups of similar receptors with overlapping responses. The next sensing state is the *bionic olfactory bulb* (BOB) composed of processing elements named for their biological counterparts, the mitral cell layer which feedforwards to the granule cell layer. Inhibitory responses are fed back from the granule layer to the mitral layer, which is also known in biology. This is another example of lateral inhibition, or the suppression of continued responses once the cell is stimulated.

The task is to discern one of seven Chinese liquors which come from different geographical locations with their own unique combination of odorants. Little is known in biology concerning how the natural olfactory systems process the spike signals for specific odor detection. The researchers here used two traditional methods for electronic nose data processing, namely linear discriminant analysis (LDA) and support vector machine (SVM), as well as backpropagation artificial neural network (BP-ANN). The latter has significant semblance to biological information processing and performed better than the other two.

5.2.4 Integrating insect olfactory receptors for biohybrid gas flow sensor [Yam20]

In this effort biology is used to create chemical sensing since the natural sensor is sensitive and selective. Insect DNA is used to synthesize olfactory receptors which are brought into an artificial cell membrane. The difficulty is getting the input gas into a soluble form for chemical detection of the artificial olfactory receptor. Microscopic slits were designed into the gas flow path and modified with hydrophobic (water-repelling) coating to create microchannels for chemical detection.

Since the odorant detection was sporadic for the given stimulants the design was scaled to monitor 16 channels. This appears to give the detection response that is desired. Biology also relies on multiple channels or opportunities for a successful chemical detection. For example, the male silkworm moth discussed earlier can detect a single molecule of the female pheromone due to antennas each having over 10,000 sensilla (each 100 microns long and 2 microns in diameter).

5.2.5 Robotic lobster chemotaxis in turbulent chemical sources [Grasso02]

The Robo-Lobster experiment is motivated by a desire for autonomy for underwater vehicles. Acoustics is primarily used and sometimes optics, but many biological species make strong use of chemo-sensing. The lobster has long antennas that sample the water chemistry for purposes such as eating, mating, spawning, and avoiding predators. A challenge to locating an odorant source is the turbulent nature of underwater chemical plumes that cause discontinuities in chemical trails; gradient descent will not work. If moving toward a food source that is detected, the lobster antennas meander back and forth in an attempt to catch samples of the odorant and the lobster adjusts its orientation and movement direction in response to what is detected. Numerous underwater chemical source detection applications exist in the scientific, environmental, commercial, and military industries.

The emphasis of the effort was more on the autonomous acquisition of the chemical source. The ability of the lobster to crawl on the bottom was simplified to an underwater wheeled robot in a fish tank. The tank measured 10m by 2m and was filled to 44 cm deep with moving seawater. A chemical source was introduced that brought odorant molecules to the robot in slow-moving turbulent patterns. The robot would move forward when the chemical was detected (when the sensor conductivity exceeding a threshold) and oriented itself so that the responses to the two artificial antennas was more balanced. Sensor responses were converted to digital values and a Motorola microcontroller programmed in C was used to implement the wheel-movement algorithm.

A robot designed to imitate a particular species and attempting to perform a task done by that species can illuminate our understanding of biology. The authors express this by suggesting “construct a robot that is competent to test a hypothesis or set of hypotheses that have been suggested by the biology and then allow the robot’s behavior to inform you of the acceptability of that hypothesis”.

Chapter 5 Questions

1. Give examples of people (and their applications) interested in a better understanding of chemo-sensory systems:

2. What are some reasons for pursuing research in bioro botics:

3. What is olfaction?

4. What is gustation?

5. Why is it difficult to differentiate olfactory and gustatory sensing systems?

6. _____The best way to distinguish between olfaction and gustation is

   a) olfaction is simple, and gustation is complex

   b) olfaction is a distance sense, and gustation is a contact sense

   c) olfaction detects chemicals in air while gustation detects chemicals in fluids

   d) there is no difference between the two

7. Photoreceptors are the first neurons in the visual processing system pathway. What are the first neurons in the olfactory system pathway?

8. The retinotopic map (vision) and somatotopic map (touch) in the brain provides a spatial map of external to stimuli in the respective systems. How is the olfactory system receptive field mapped in the brain?

9. Taste buds provide neuronal inputs to what type of cells?

10. What do hygro-receptors detect?

11. What part of the insect have we always found hygro-receptors?

12. Differentiate between a semiochemical and a pheromone.

13. What is anemotaxis?

14. What is rheotaxis?

15. What are some application areas for successful lobster chemotaxis research?

16. What were some of the significant results from MIT's robotic implementation of lobster chemotaxis in Robo-Lobster?

Chapter 5 References

[Consi94] Consi, T. R., et al., "AUV guidance with chemical signals", in *Proceedings of the IEEE Symposium on Autonomous Underwater Vehicle Technology*, p. 451, 1994.

[Grasso02] Grasso, F., "Flow and chemo-sense for robot and lobster guidance in tracking chemical sources in turbulence", Chapter 27 in Ayers, J., Davis, J., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, ISBN 0-262-01193, 2002.

[Liu18] Liu, Y., et al., "Using spike-based bio-inspired olfactory model for data processing in electronic noses", *IEEE Sensors Journal*, Vol. 18, No. 2, 2018.

[Persaud82] Persaud, K. and Dodd, G., "Analysis of discrimination mechanisms in the mammalian olfactory system using a model nose", *Nature*, Vol. 299, No. 5881, pp. 352–355, Sept. 1982.

[Smith08] Smith, C. U. M., *Biology of Sensory Systems*, 2nd ed., John Wiley and Sons, ISBN: 978-0-470-51862-5, 2008.

[White02] White, J., et al., "Using biology to guide development of an artificial olfactory system", Chapter 5 in Ayers, J., Davis, J., and Rudolph, A., Eds., *Neurotechnology for Biomimetic Robots*, MIT Press, ISBN: 0-262-01193, 2002.

[Yam20] Yamada, T., et al., "Odorant sensor using olfactory receptor reconstituted in a lipid bylayer membrane with gas flow system", *IEEE MEMS 2020*, Vancouver, CA, ISBN: 978-1-7281-3581-6, 2020.

Appendix A – Example Literature Review Assignment

The following is a representative literature review assignment. The course website was hosted by CANVAS with an active discussion board.

Use the CANVAS discussion board to claim the paper you would like to review – make sure it has not already been claimed. If you would like to consider other papers, that is fine, but email instructor a copy to approve before you do the review. It must be recent and directly relevant to imaging systems whose designs (or novel parts of it) are inspired by natural imaging systems (vision).

The FSU Libraries are very useful for finding additional publications. If interested:

> Search "FSU Libraries" and then "Find a Database". Under the option to "Search our A-Z list of Databases" pull down "I" and find IEEExplore. At this point you may have to log in to the FSU Portal. Once in IEEE Xplore, you can search on "biologically-inspired", "bio-inspired imagers", or some other related term. On the left under "year" refine your search by moving the range to the most recent year-and-a-half (2018 to 2019), and the click on "Apply refinements".

Once you have selected one of the provided papers (or have approval for a different paper) enter the citation (at least the author, title and year) on the course discussion board and confirm no one else has selected that paper. Read the paper and study it well enough to discuss it in class. You are not required to understand all derivations, equations, etc. but should be able to answer the following questions in a Word file. Your answers do not have to be long but should be in your own words and very clear and accurate; there is not a requirement on length. Use sentences and not phrases or bullet points. While presenting in class pull up your Word document (avoid power-point, etc.) and you may pull up the paper you are reviewing as well; feel free to go back and forth between your written answers and the paper. You may refer to the figures, tables, diagrams, or anything else in the paper, but your Word file should answer the questions without the reader having to refer to the paper.

Turn in a Word file or PDF file that answers the questions in this format like the attached example on the next page:

Reviewed by{*your name*}

Paper citation: {*citation*}

What is the problem to be addressed or solved?

What is the natural paradigm being considered?

What has already been done?

How is this approach different?

What accomplishment is claimed?

What do they plan to do next?

{Graduate students} Discuss at least one (or more) of the mathematical derivations or equations in the paper. If there are no derivations discussed, then choose a paper which does.

Paper Abstract (pasted): {*paste paper abstract here*}

Post your Word file (or PDF) in the course Assignment folder. If you selected a paper not yet posted, post a copy of it as well with your file.

Example Paper Review

Reviewed by Dr. Brooks

Paper citation: H. Wu, K. Zou, T. Zhang, A. Borst, K. Kuhnlenz, Insect-inspired high-speed motion vision system for robot control, Biological Cybernetics, 106:453-463, 2012).

What is the problem to be addressed or solved?

Improve the accuracy of velocity estimation in the Hassenstein-Reichardt Elementary Motion Detection (HR-EMD) model. Velocity estimation of the objects in an image is integral to visual perception and will be a necessity for robot control systems performing autonomous navigation and collaboration with other agents (robots). Motion estimation using conventional imaging technology is slow.

What is the natural paradigm being considered?

Motion detection at the neuronal level in insect vision, specifically the well-known HR-EMD model.

What has already been done?

The basic insect-vision-inspired HR-EMD model is well established. It has been used to address aircraft guidance (collision-avoidance, gorge-following, and landing) and demonstrated in robotic platforms. It has been implemented in VLSI for collision detection and implemented in FPGAs for optic flow detection and motion estimation. It has been applied for course stabilization and altitude control of a blimp-based unmanned aerial vehicle.

How is this approach different?

The former applications of the HR-EMD are based on a more qualitative motion detection and not a quantitative motion velocity estimation. Here the authors are using image pattern statistics (brightness, contrast, and a spatial PSD estimation) combined with the HR-EMD output and by a look-up table estimating the velocity of motion instead of simply motion. Here also, as with former efforts of the authors, a conventional temporal low-pass-filter is used as the delay element in the HR-EMD.

What accomplishment is claimed?

The average EMD response of the entire image was used for closed-loop yaw-angle control system of a robotic manipulator arm. They demonstrated yaw control using a piece-wise linear motion input and an arbitrary motion input.

What do they plan to do next?

They plan to extend to demonstrate motion estimation of 3D objects in a receptive field.

Paper Abstract (pasted): **The mechanism for motion detection in a fly's vision system, known as the Reichardt correlator, suffers from a main shortcoming as a velocity estimator: low accuracy. To enable accurate velocity estimation, responses of the Reichardt correlator to image sequences are analyzed in this paper. An elaborated model with additional preprocessing modules is proposed. The relative error of velocity estimation is significantly reduced by establishing a real-time response velocity lookup table based on the power spectrum analysis of the input signal. By exploiting the improved velocity estimation accuracy and the simple structure of the Reichardt correlator, a high-speed vision system of 1 kHz is designed and applied for robot yaw-angle control in real-time experiments. The experimental results demonstrate the potential and feasibility of applying insect-inspired motion detection to robot control.**