

1991

Optical photoreceptors for CMOS neural circuitry.

Naveen. Bewtra
University of Windsor

Follow this and additional works at: <http://scholar.uwindsor.ca/etd>

Recommended Citation

Bewtra, Naveen., "Optical photoreceptors for CMOS neural circuitry." (1991). *Electronic Theses and Dissertations*. Paper 2938.

This online database contains the full-text of PhD dissertations and Masters' theses of University of Windsor students from 1954 forward. These documents are made available for personal study and research purposes only, in accordance with the Canadian Copyright Act and the Creative Commons license—CC BY-NC-ND (Attribution, Non-Commercial, No Derivative Works). Under this license, works must always be attributed to the copyright holder (original author), cannot be used for any commercial purposes, and may not be altered. Any other use would require the permission of the copyright holder. Students may inquire about withdrawing their dissertation and/or thesis from this database. For additional inquiries, please contact the repository administrator via email (scholarship@uwindsor.ca) or by telephone at 519-253-3000ext. 3208.



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser à désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure.

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.



Optical Photoreceptors for CMOS Neural Circuitry

by

Naveen Bewtra

A Thesis

Submitted to the Faculty of Graduate Studies through the
Department of Electrical Engineering in Partial Fulfillment
of the Requirements for the Degree of
Master of Applied Science at the
University of Windsor

Windsor, Ontario
March 1991



National Library
of Canada

Bibliothèque nationale
du Canada

Canadian Theses Service Service des thèses canadiennes

Ottawa, Canada
K1A 0N4

The author has granted an irrevocable non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of his/her thesis by any means and in any form or format, making this thesis available to interested persons.

The author retains ownership of the copyright in his/her thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without his/her permission.

L'auteur a accordé une licence irrévocable et non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de sa thèse de quelque manière et sous quelque forme que ce soit pour mettre des exemplaires de cette thèse à la disposition des personnes intéressées.

L'auteur conserve la propriété du droit d'auteur qui protège sa thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

ISBN 0-315-72812-4

Canada

Naveen Bewtra 1990

All Rights Reserved

To my mother and father

Abstract

The work described in this thesis relates to the design of photoreceptors fabricated in Northern Telecom's 3μ CMOS process. The photoreceptor consists of a vertical BJT whose output current is fed to four diode connected MOSFETs. Results of the experiments used to characterize the device both optically and electronically are given. Shortcomings of the device are identified and a modified photoreceptive device has been designed.

These photoreceptive devices will be used to provide optical inputs to a neural network. By using optical connections to provide inputs to the chip, data can be sent directly to individual locations on the chip from above. This approach may eventually be used to solve the problems of an analog storage device and of pin-limitations usually associated with the VLSI implementation of neural networks. An architecture for making optimal use of the photoreceptors characteristics is given. This current mode architecture uses the current from the photoreceptor as the input to the network as well as the weights.

Acknowledgements

I would like to take this opportunity to thank the guidance and support provided by Dr. W.C. Miller and Dr. G.A. Jullien during my work. Their patience and understanding will always be remembered. I would also like to thank Dr. S. Bandyopadhyay for his role as my supervisory committee member. A special thanks must go to Mr. Ajay Chandna, an excellent researcher, but more importantly an excellent friend.

TABLE OF CONTENTS

Abstract	v
Acknowledgements	vi
Table of Contents	vii
List of Figures	ix

CHAPTER 1

INTRODUCTION TO NEURAL NETWORKS

1.1 Neural Information Processing	1
1.2 Neurons	3
1.3 Artificial Neural Networks	6
1.4 VLSI Neural Network Implementations	11
1.5 Thesis Organization	15

CHAPTER 2

REALIZATION OF PHOTSENSITIVE DEVICES IN CMOS TECHNOLOGY

2.1 Introduction to Theoretical Considerations	17
2.1.1 Conceptual Design	17
2.1.2 Operation of the BJT and Photodiode	19
2.1.2.1 Operation of the BJT in the Active Mode	19
2.1.2.2 Current Gain in the Transistor	22
2.1.2.3 High Frequency Bipolar Model Parameters	24
2.1.2.4 Operation of the Photodiode	26
2.1.2.5 Response time of the Photodiode	30
2.1.3 Phototransistor Operation	33
2.1.4 Conclusions	37
2.2 VLSI Implementation of the Photoreceptor	38
2.2.1 Introduction to Fabrication of Photoreceptive Devices in CMOS	38
2.2.2 VLSI Implementation of the Photoreceptor	38
2.2.3 Practical Design Considerations	42
2.3 Conclusions	45

CHAPTER 3

CHARACTERIZING THE PHOTORECEPTIVE DEVICE

3.1 Determination of Output Voltage Range	46
3.2 Removal of Optical Crosstalk by P-Well	48
3.3 Characterizing the Vertical BJT	50

3.4. Logarithmic Response of the Photoreceptor	53
3.4.1 Static Characteristics	54
3.4.2. Experimental Procedure	54
3.4.3 Experimental Results	55
3.4.4 Conclusion	56
3.5 Transient Response	57
3.6 Spectral Response	59
3.7 Determination of the Early Voltage	61
3.7.1 Experimental Setup	61
3.7.2 Cause of the Early Voltage	62
3.7.3 Effects of the Early Voltage	63
3.8. Various Photoreceptive Configurations	64
3.8.1 Active Feedback Photoreceptor	64
3.8.1.2 Static Response Properties	65
3.8.1.3 Comparison with the Logarithmic Receptor	68
3.8.1.4 Conclusions	69
3.8.2 Optical Detector using MOSFETS	70
3.8.2.1 Operation of the Device	71
3.8.2.2 Summary	75
3.9 Conclusions	75

CHAPTER 4

AN OPTICALLY COUPLED NEURON FOR NEURAL NETWORKS

4.1 Introduction	77
4.2 Approximate Multiplication by Correlation	81
4.3 Thresholding the Weighted Input	86
4.4 Conclusions	92

CHAPTER 5

SUMMARY AND CONCLUSIONS

5.1 Summary and Conclusions	93
5.2 Future Work	94

REFERENCES	96
------------	----

APPENDIX I	Plot of the Photoreceptive Device	99
APPENDIX II	Theoretical Basis of the Liquid Crystal Device	100
APPENDIX III	Spice Deck and Model Card for Photoreceptor	123
APPENDIX IV	Circuit for Triggering Laser	125

Vita Auctoris	128
---------------	-----

List of Figures

Figure	Title	Page
1.1a	Block Diagram of Biological Neuron	4
1.1b	Block Diagram of Artificial Neuron	5
1.2	Block Diagram of Feed Forward Network	7
1.3	Block Diagram of Feedback Network	8
1.4	Various Thresholding Functions	9
1.5	Block Diagram of Four Neuron Network	10
2.0	Conceptual Schematic of Photoreceptor	18
2.1	Current Flow in a BJT	20
2.2	Minority Charge Distribution of NPN	21
2.3	Hybrid Pi Model showing Capacitance	24
2.4	Energy Diagram for a Photodiode with Absorption	26
2.5	Photodiode Equivalent Circuit	27
2.6	Optical Absorption Coefficient for Si	29
2.7	Photodiode Response Time	33
2.8	Cross Sectional View of Phototransistor	34
2.9	One Dimensional Model of Phototransistor	34
2.10	Circuit Schematic and Cross Sectional Layout	39
2.11	Simulated Photoreceptor Performance	41
2.12	Expected Experimental Results	44
3.1	Actual Photoreceptive Results	47
3.2	Effect of P-Well Modification	49
3.3	Collector Current vs VCE	51
3.4	Beta as a Function of Emitter Current	52
3.5	Vout vs Filter ND	56
3.6	Time Response of the Photoreceptor	58
3.7	Spectral Response of the Photoreceptor	59
3.8	Early Voltage Characteristics of BJT	61
3.9	Minority Charge Concentration	62
3.10	Active Feedback Device	66
3.11	Modified Active Feedback Device	67
3.12	Simulated Results of Active Feedback Device	68
3.13	Simulated Results showing Faster Response	70
3.14	Electron Hole Pairs in MOSFET	72
3.15	Ids vs Vds for MOSFET with and without Light	74
4.1	Artificial Neuron Block Diagram	78
4.2	Block Diagram of the Current Mode Neural Network	79
4.3	Correlator Circuit and Spice Simulations	82
4.4	Block Diagram showing use of Four Quadrants	86

4.5	Four Quadrant Correlator	87
4.6	Current Mode Operational Amplifier	88
4.7	Current Operational Amplifier Simulation	90
4.8	Entire Differential Current Mode Neural Network	91
4.9	Simulation Results of Network	92

I. Introduction to Neural Networks

An artificial neural network is a computing system that combines a network of highly interconnected processing elements mimicking the physiology of the brain. The processing elements, which contain mathematical algorithms, perform information processing through their state responses to stimuli.

Early attempts to create artificial neural systems failed simply because no workable technology existed for realizing systems of the required complexity. It was for that reason that until recently there existed little interest in studying neural networks. However, the rapidly developing technology of very large scale integrated (VLSI) circuits has provided a medium in which it is presently possible to fabricate tens of millions of devices interconnected on a single silicon wafer. This technology has resulted in a renewed thrust in researching the neural network and its properties, as well as its implementation in Si. This chapter examines some of those properties and how far the VLSI implementation of neural networks has come.

1.1 Neural Information Processing

Digital computers are ideal for algorithmically derived problems, sorting, formal logic and probability calculations, whereas the human brain tends to do those tasks slowly and poorly. The brain is however adept at very complex pattern recognition.

generalization, intuition, problem finding, and language. It seems that the architecture and function of the brain is better adapted to these "characteristically biological" tasks than are digital computers. Neural information processing is a type of processing that attempts to copy the functionality of the human brain in solving taxing pattern-processing problems.

This type of processing can be divided into two separate areas: *neural science* and *neural computing*. The first division, neural science, deals mainly with the study of brain functions and the modeling of these functions. Both neurophysiology and the biophysics of biological neural networks are investigated in order to determine the large number of complex processes involved in the activity of neurons [Sh87].

The second division of neural information processing is neural computing. Neural computing researches artificial neural networks, neurocomputers, and neural programming environments and attempts to apply them to a wide range of pattern-processing applications. Pattern processing includes

- image processing
- speech processing
- inexact knowledge processing
- natural language processing
- sensor processing
- planning

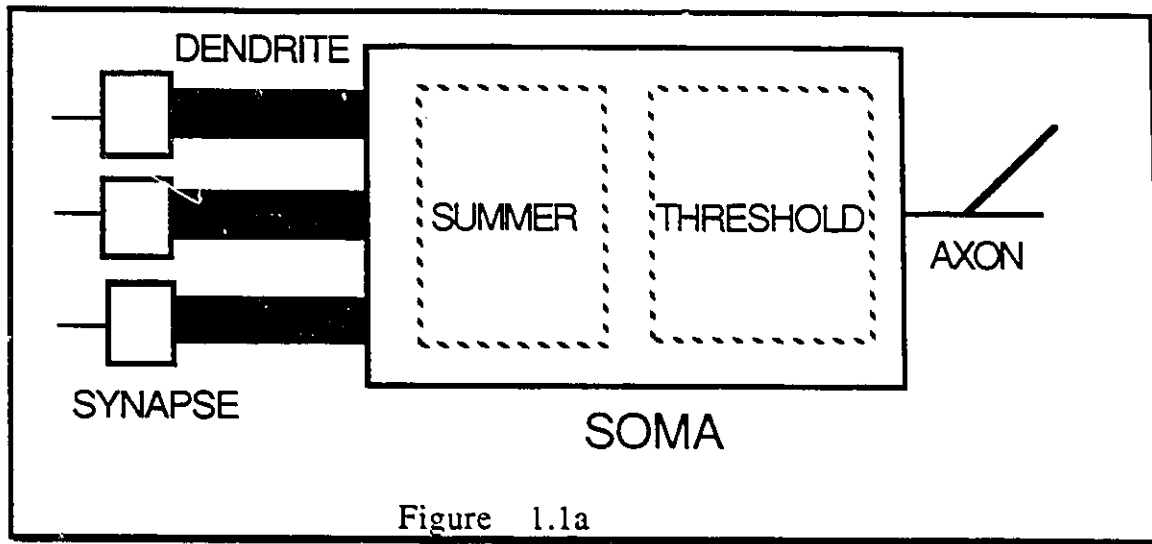
- forecasting
- optimization

For these problems, artificial neural networks exhibit various properties which are similar to the functions of a brain, including association, generalization, learning and flexibility [DeWi87].

1.2 Neurons

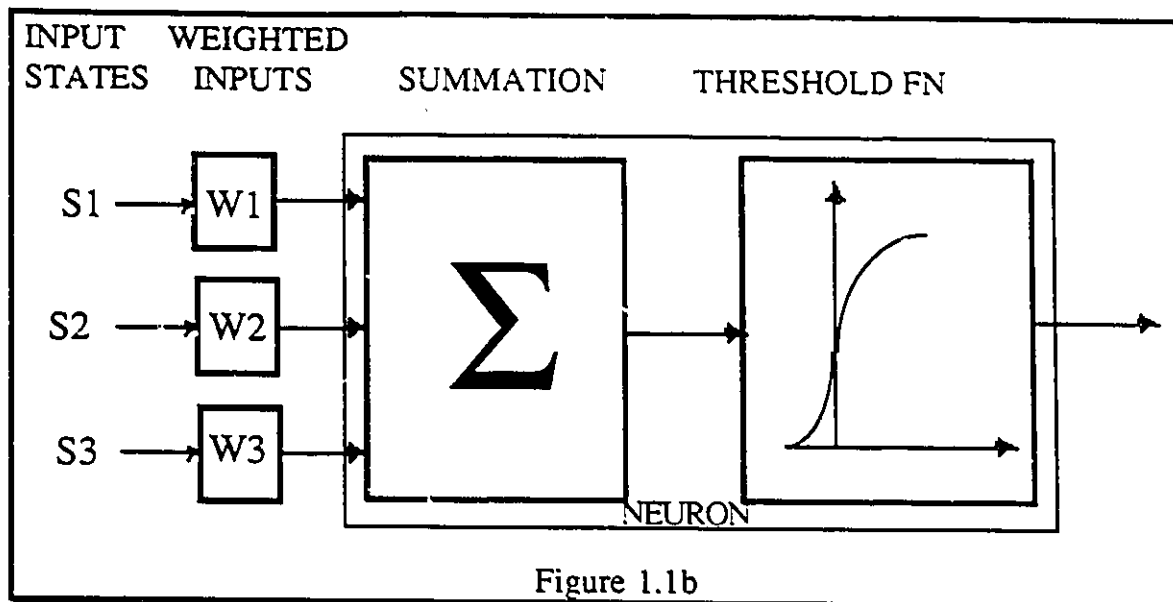
Certain researchers feel that neural network circuits can more easily be designed if one has a thorough understanding of the brain and the way its neuronal systems operate [Mead89]. Yet the brain is extremely intricate and not easily understood. However, some aspects of the human brain have been mimicked, as shown in Figure 1.1a, which illustrates in very simplified terms, the brain structure and the basic principles that nerve cells use for information processing.

Figure 1.1a represents the biological neuron. It consists of a cell body, or *soma* (which has complex extensions that serve as inputs, also called dendrites), and the output channel of cell otherwise known as the *axon*. Electrical signals are carried by the axon to other cells. The axon then connects to the dendrites of the cells through specialized contacts called synapses. These synapses are able to change, positively or negatively, the axon potential. The neuron's function in brain activity is determined by its shape and interconnection characteristics.



In the past it was thought that the neuron performed a simple threshold function: weighted input signals were summed and if the result exceeded a particular threshold, a signal would be emitted from the neuron. Recently though it was discovered that the dendrites and synapses are actually more involved than previously thought. It is believed that some sort of computation takes place outside of the neuron body.

The complexity of the brain is staggering. It consists of approximately 10^{11} neurons, with each neuron connected to approximately 10^4 other neurons. A neuron fires by sending an electrical impulse that leaves its cell body and reaches the next neuron through the synaptic junction, and if the next neuron has enough energy at its inputs, it will also fire. Although the firing rates of the neurons are low, the brain is able to solve difficult problems such as vision and association[Mead89].



An artificial neuron, shown in Figure 1.1b, is comprised of three distinct parts. These parts include the weighted inputs connections, the summation function, and a threshold device, which depends on the states (s_i) of its input elements and the connection strengths, or weights (W).

In an artificial neuron network, neurons are generally configured in regular and highly interconnected topologies. In the Hopfield model, a single layer of artificial neurons form the topology. The output from each neuron feeds back to all of its neighbours. Other models such as Boltzmann Machine consist of one or more layers between the input and output neurons. Still other models like the Self-Organizing Map the network connects a vector of input neurons to a two-dimensional grid of output neurons [TrPaVe89].

The correspondence between the brain and an artificial neural network is quite significant. The major difference between the two is the switching speed. A biological neuron switches on the order of 1 millisecond where a transistor can switch in 1 nanosecond. The difference in switching speeds illustrates the brains use of parallelism as it takes less than 100 steps to process an intricate task such as vision in a fraction of a second [Mead89].

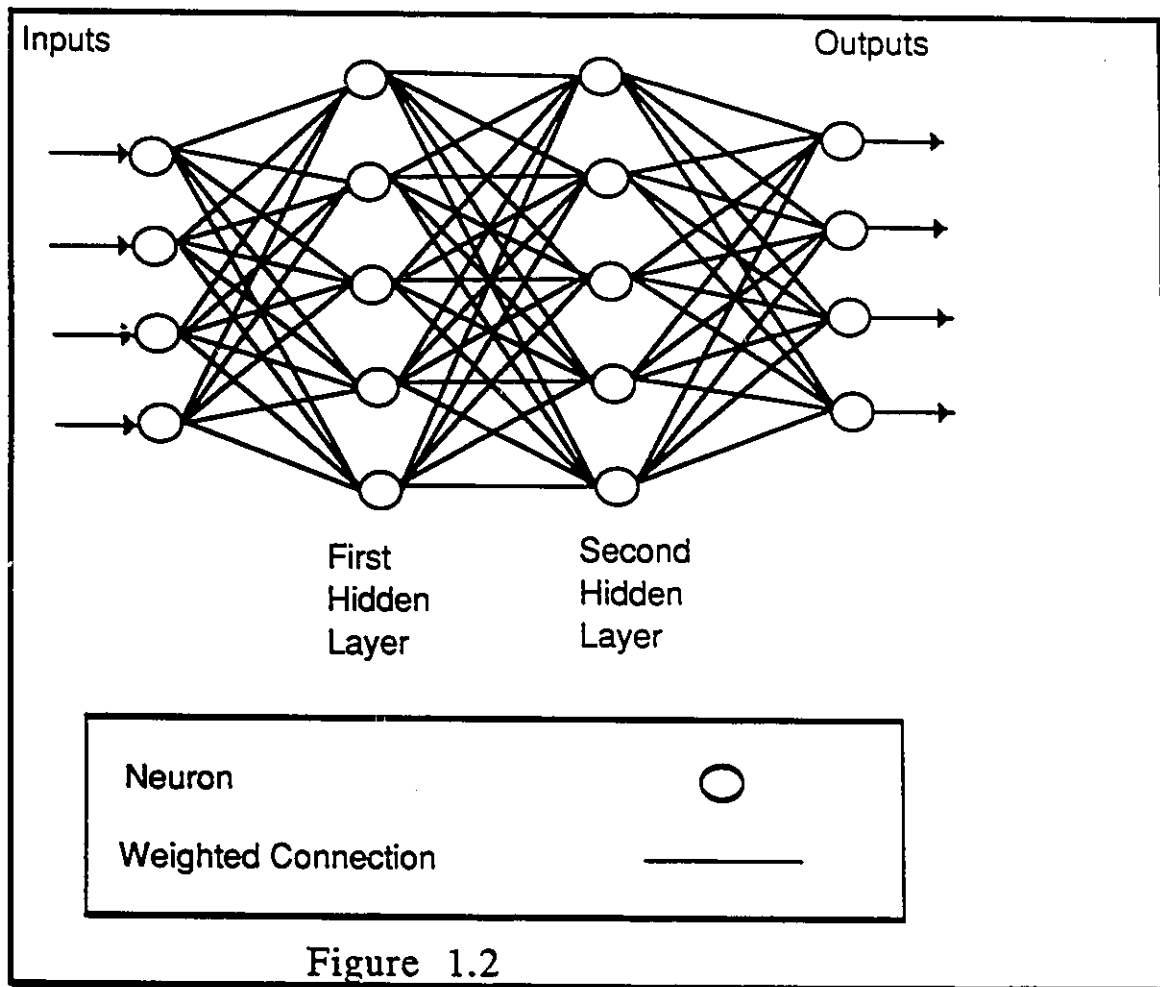
In summary, it can be stated with some degree of confidence that the biological neuron is far more complex than the artificial neuron. During the normal functioning of biological neurons, information processing can involve many complex processes. Artificial neural research, on the other hand, started with a simple threshold model and added complexity only when new knowledge is obtained and the old model was improved. However, this approach has been the cornerstone to the advances made in the neural network field.

1.3 Artificial Neural Networks

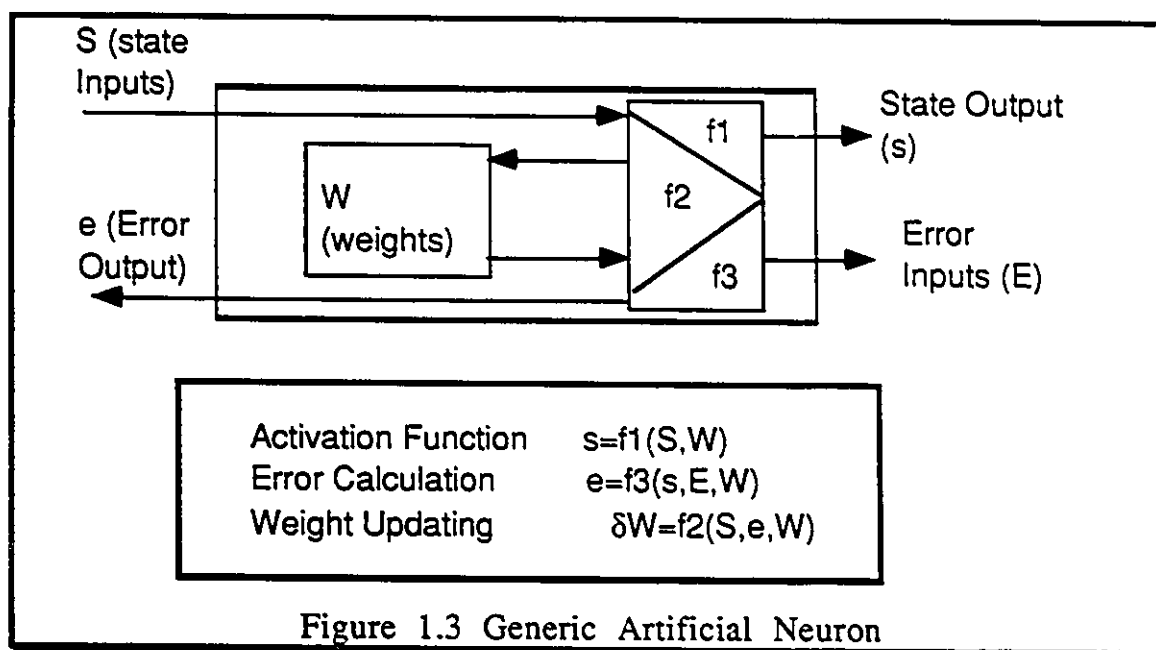
Artificial neural networks are directed graphs represented by weighted interconnections between neurons. The data of the network is represented by the weights and the states of the neurons. By adjusting the weights of the connections between the neurons, a neural network learns to recognize patterns. Neural networks possess a few key properties. The first of these key characteristics is network topology, which is the neuron interconnection pattern. Neurons are usually arranged into structures known as layers. The

first models were comprised of a single layer and each network input was connected to all neurons. In such a model, the information would flow from input to output with no feedback paths, thereby given the name feed-forward networks [TrPaVe89].

More recent models have expanded the number of feed-forward layers, as shown in Figure 1.2. In these models, the middle layers are hidden from external input and output, but they receive weighted input.



When the main variations of learning and recall procedures that are found in different classes of incorporated a generic artificial neuron results. As illustrated in Figure 1.3, this type of neuron consists of three specific functions (f1 [the activation function], f2 [the weight-updating function], f3 [the error-calculation function]), a table of weights (W), and sets of input (S, E) and outputs (s, e). A neuron receives the input states S from the previous layer's neurons, and forwards its state output s. Errors are processed in a similar manner providing feedback for the network.



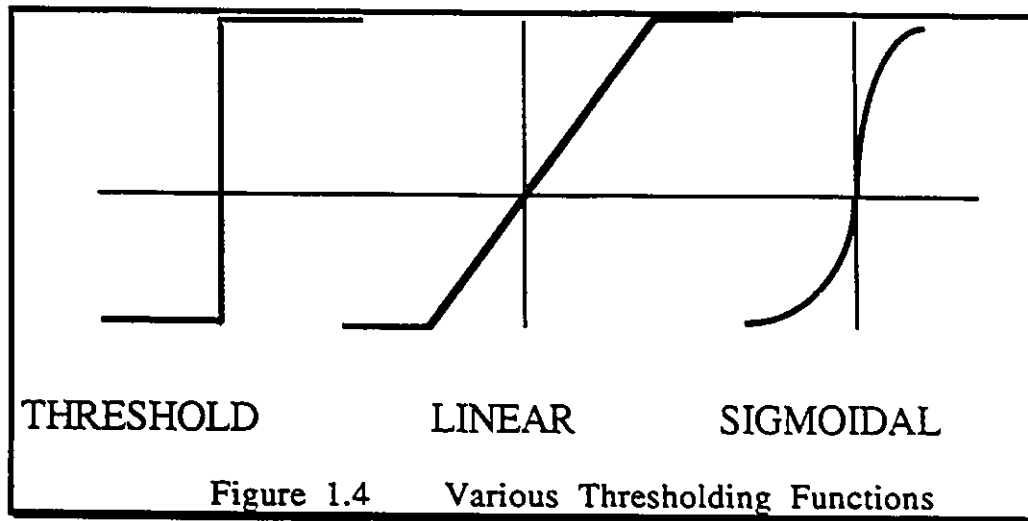
Another key property of neural networks is the recall procedure. This procedure is specified by the activation function f1, which comprises the propagation rule net and the threshold function

$$\text{net} = F(S,W) \text{ and } s = T(\text{net}) \quad (1.1)$$

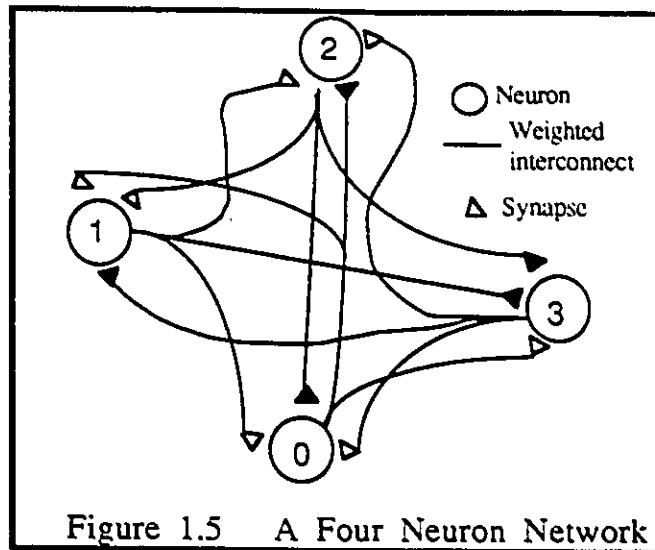
The propagation rule calculates the weighted sum of the input, modified by an offset q that defines the neuron bias:

$$\text{net} = \sum S * W - \theta \quad (1.2)$$

The nonlinear function of the propagation rule calculates the neuron state. Hard limiter, sigmoid, and pseudolinear, represent some common threshold functions (Figure 1.4) [RuMc86].



The final property of neural networks is the input values. The input values are generally characterized by the range they take on, namely, continuous or discrete valued input [MuSm88].



An overview of how these properties work together in the network will now be discussed. Figure 1.5 shows a network consisting of 4 neurons. A neuron can be thought of as a state machine that is normally ON or OFF. It may take on intermediate states between these extremes in a smooth transition. As discussed in the previous section, a biological neuron signal is in the ON position by firing an electrical pulse along its output or axon. Also it may be in the OFF position by ceasing to fire pulses. Each axon may separate to form connections to all other neurons in the network, in the most complicated case to each of the other neurons in the network. These connections are made through synapses (as discussed earlier), which are represented by triangles in Figure 1.5. The function of the synapse is to gate the signal from the transmitting neuron such that it may be seen as an excitatory (filled triangle) or inhibitory (unfilled triangle) signal to the receiving neuron. The excitatory input tends to turn the receiving neuron on, whereas the inhibitory input seems to turn the receiving neuron off. If neuron 3

in Figure 1.5 receives an output from neuron 1 as an input, gated by a factor W_{31} ; W_{31} may be positive, negative, or zero (it is shown as excitatory in Figure 1.5). The synapses have therefore the effect of weighting the output response of any connected neuron to its inputs, and the total weighted sum changes the level of activity of the neuron. [MuSm88]

The network shown in Figure 1.5 can therefore have an infinite number of possible states, corresponding to all combinations of the individual neural states (ON, OFF, or some intermediate value). The set of synaptic weights determine the states, and represents the information learned in the system. Clearly then, learning is a process where the synaptic weights are changed to add to the network's store of knowledge. If the interconnect weights between neurons 0 and 2 and those between neurons 1 and 3 are set as large and positive (excitatory), and all other interconnect weights are negative. Therefore the stable states of the network will clearly be those with neurons 0 and 2 ON and 1 and 3 OFF. If the network is forced into some other state, it will move toward one or the other of these two preferred states. It is this dynamic behavior that is at the heart of neural computation.

1.5 VLSI Neural Network Implementations

Neural network implementations fall into two broad categories, digital and analog. Currently research is being done in both areas. The best way to present the information is in a comparison between

the two distinct approaches, thus showing their strengths and weakness [MuSm88].

A. Digital Neural Networks

The strengths of a digital approach can be summarized as:

ADVANTAGES

- . design techniques are advanced, and well automated
- . noise immunity is high
- . computational speed can be very high
- . learning networks (i.e. those with programmable weights) can be implemented readily

However, for neural networks, there are several unattractive features:

DISADVANTAGES

- . digital circuits of this complexity are synchronous, whereas real neural nets are asynchronous
- . All states, activities, etc in a digital neural network are quantized
- . Digital multipliers, essential to the neural weighting function, occupy large silicon area

B. Analog Neural Networks

The benefits of analog networks are more subtle :

ADVANTAGES

- . asynchronous behavior is automatic
- . smooth neural activation is automatic
- . circuit elements can be small

However, for neural networks, there are several unattractive features:

DISADVANTAGES

- . noise immunity is low
- . arbitrary high precision is not possible

In both digital and analog realizations, the largest computational load is incurred by the weighted summation

$$\sum_{j=0}^{j=n-1} T_{ij} V_j \quad (1.3)$$

where T_{ij} represents the synaptic weight between neuron i and j

V_j represents the input to neuron j

Most effort in implementation of VLSI networks is concentrated on forming this sum, which contains (potentially) n^2 terms.

In both analog and digital realizations, a problem with VLSI neural implementations is apparent: the stored weights must be made available to the multiplication circuitry. In an architecture

where a small number of multipliers is used, this requires accurate communication and fast memory access.

The same problem shows up in analog networks, although it is relatively small compared to the problem of storing the weights on chip. As a result, a great deal of analog VLSI implementations are nonprogrammable, and therefore have fixed functionality. Subthreshold MOS device characteristics allow much smaller power consumption on the chip and have been used to mimic the nonlinearities of biological neuron behavior. The Subthreshold region of operation has been used in implementing Hopfield nets [SiEmMe86], associative memories [SiEmMe85], visual processing functions [Mead89], and auditory processing [Mead89]. Another major research group uses electron beam programmable resistive interconnects to represent synaptic weights between more conventional operational amplifier neurons. This approach has led to the concept of a "floating gate technology" which uses Ultra-Violet light to program the weights and store them on chip. This work has led to an associative memory device using on chip memory to store weights and analog circuitry to perform the arithmetic functions [GrVe89]. An significant development is the use of charge coupled devices (CCD)/MNOS (metal nitride oxide silicon) technology to store analog weights, and thus keep the entire neural network analog and yet programmable [HiGo88]. The disadvantage that has been a source of concern is the time required to program the weights, as well as the amount of time the weights can stay on chip. It is felt that eventually this will be the direction that will be used to

implement fully analog memory on the chip even if unusual technology is required.

This thesis deals with the development of photoreceptive devices for use in optically coupled analog neural networks. Light of varying intensities (weights) is focused on photoreceptive devices, thus moving the problem of storing weights off chip to provide a mechanism for optically programming the large number of weights usually associated with neural networks.

1.6 Thesis Organization

The thesis is broken down into four main chapters as can be seen by the table of contents.

The second chapter "Realization of Photosensitive Devices in CMOS Technology" contains two very important sections. The first deals with the theoretical background on the operation of the bipolar junction transistor (BJT) and the photodiode. Once this is completed, a first order model of the phototransistor, containing a BJT and a photodiode is presented. This section explains the operation of the phototransistor as photons strike the surface of the device. The second major part of this chapter deals with the practical implementation of the phototransistor in Northern Telecoms 3u CMOS process. The construction of the BJT in this technology is given, as well as the effect of the diode-connected MOSFETS. Also included is a discussion of some practical design considerations that are implemented in order to enhance the robustness of the device. Using

the theoretical background presented, the performance of the device is predicted.

The third chapter "Characterizing the Photoreceptor" provides the experimental characterization of the photoreceptor implemented in the Northern Telecom process. Included is a discussion of experimental results including current gain, optical dynamic range, and time response. These results are vital for the application of these devices as optical inputs to VLSI circuitry. Also included in this chapter are various implementations used to enhance the performance of the device. The circuitry will be given, as well as a discussion of expected results.

The fourth chapter "An Optically Coupled Differential Current Mode Neuron for Neural Networks" presents the implementation of the photoreceptive device in a neural network especially designed to make use of the possibility of optical programming. This section will give a description of how this current mode circuitry makes optimal use of the photoreceptive device, allowing light to bring in the inputs and weights to the neural network.

The appendices include plots of the photoreceptive device, a discussion of BJT operation, and an investigation into the fabrication of liquid crystal devices. A liquid crystal device was fabricated in order to determine the possibility of implementing a optically input and output neural network.

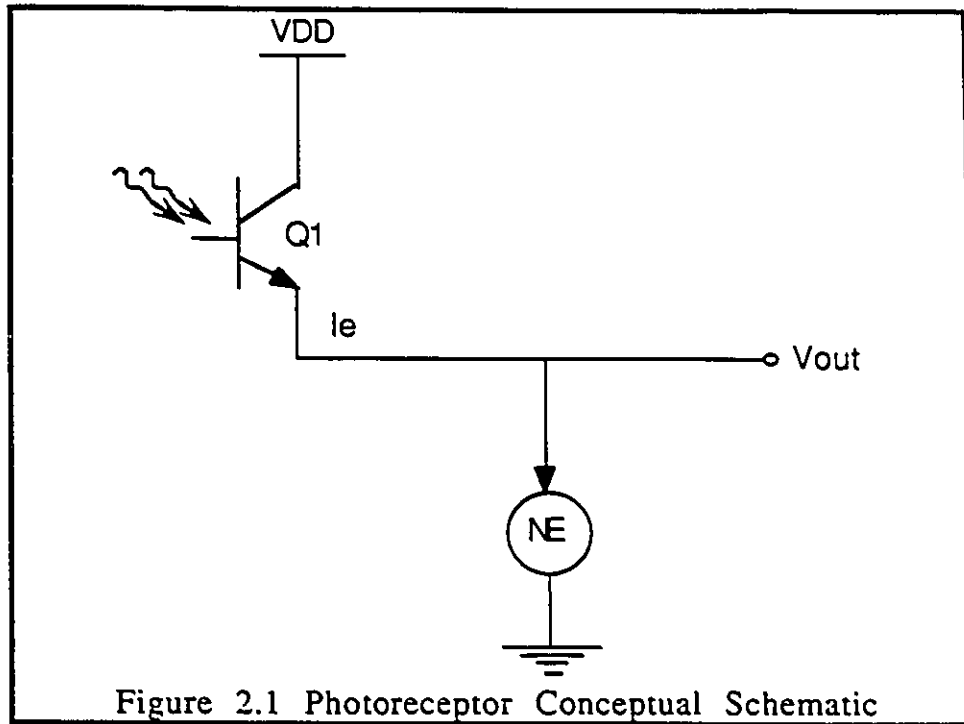
II. Realization of Photosensitive Devices in CMOS Technology

2.1 Introduction to Theoretical Considerations

In the following sections, a description of the theoretical basis for the implementation of the photoreceptor in Northern Telecoms 3 μ CMOS process is discussed. The conceptual design will be discussed in order to provide the reader with an overview of what features a suitable photosensitive device should contain. This is followed by a discussion of current flow in a BJT. Also discussed is the modelling of the phototransistor using a photodiode and a BJT. The photodiode operation is explained from a photonical and electrical point of view. Finally the phototransistor operation is discussed, in a comprehensive manner.

2.1.1 Conceptual Design

For a transducer operating over many orders of magnitude, the most appropriate transfer characteristic is logarithmic. In the figure below the circuit for generating an output voltage proportional to the logarithm of the intensity is given [Mead87]. The current, $I_e = \beta \times I_g$, where β is the transistor current gain and I_g represents the base current which results from minority carriers generated by the light. A non-linear element (NE) converts this current into an output voltage that provides a logarithmic response.



The device characteristics can be expressed as:

$$I = e \alpha V_{out} \quad (2.1)$$

the above equation can be reduced to

$$V_{out} = \ln (I)/\alpha \quad (2.2)$$

Any particular implementation of the receptor requires a realization of both the amplifier and the non-linear element [Mead87]. One implementation is shown above. The primary detector is the open base BJT phototransistor. The amplification is the current gain β value. The amplified current is turned into a voltage by the load devices: the chain of diode connected MOSFETS.

The nonlinear element should produce a voltage proportional to the exponential of the output current in such a way that the desired range of light intensity corresponds to a range of V_{out} between VDD and ground. The best exponential device we have is a MOS transistor operating in sub-threshold. In this region of operation, drain currents are exponential with gate voltages over many decades.

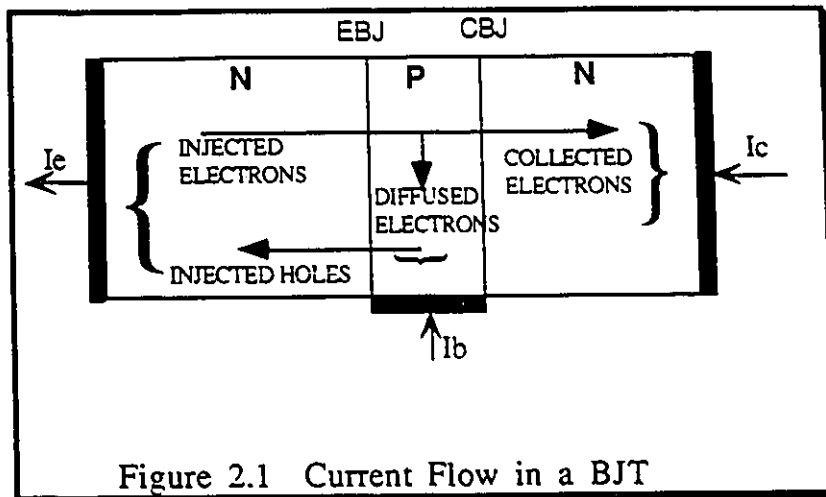
The presented photoreceptor schematic provides meaningful outputs for over several orders of magnitude.

2.1.2 Operation of the BJT as a Phototransistor

In this section, the operation of the phototransistor will be examined. This will allow a greater understanding of the results discussed in the following chapter. A brief description of the operation of the bipolar junction transistor (BJT) will be followed by a discussion of optical absorption in semiconductors. An explanation of the photodiode will be included, since this device appears in the final section on phototransistor models and operation.

2.1.2.1 Operation of the BJT in the Active Mode

It is important to discuss the physical operation of the BJT operating in the active region of operation, since this is the region the phototransistor will be operating in. In order to operate in the active mode, the two pn junctions must be biased appropriately. The base emitter junction should be forward biased, while the base collector junction should be reversed bias [SeSm87].

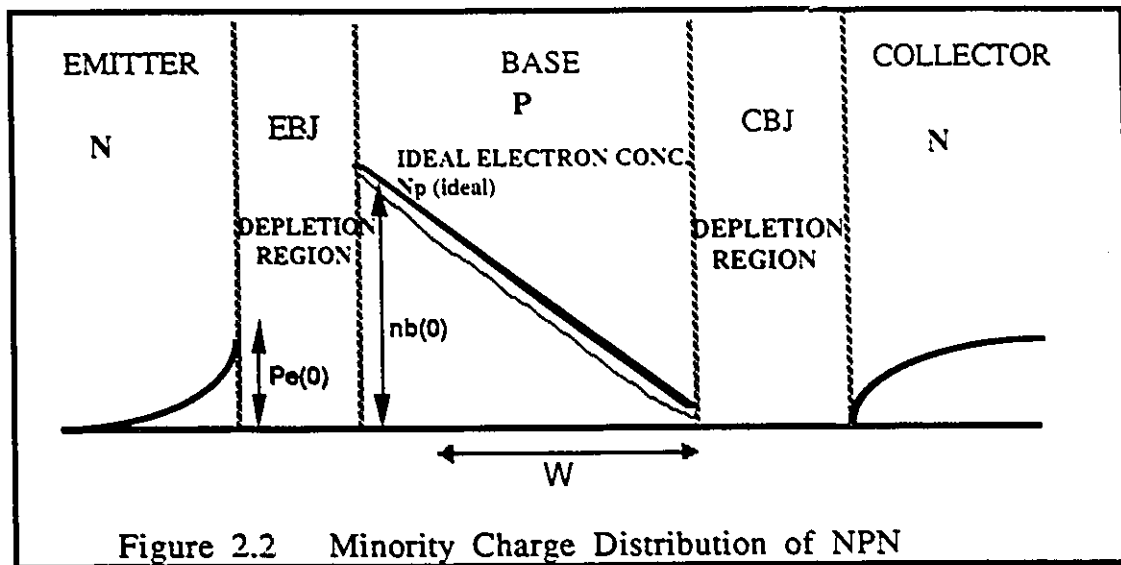


Since diffusion current is the major component of current, this is the current that will be examined more closely. However, drift current due to thermally generated carriers are also present, although these are less significant. The forward biasing of the emitter base junction results in a current flowing through the pn junction. This current consists of holes injected from the base (where they are majority carriers) to the emitter (where they are minority). Also, electrons from the emitter are injected into the base, where they become minority carriers. This component should be higher since it is these electrons which will be collected at the collector. It is for this reason that the emitter is more heavily doped than the base.

As shown in the Figure 2.1, the current that flows out of the emitter constitutes the emitter current I_e . The direction of I_e is "out" of the emitter lead which is in the direction of the hole current, and opposite to the electron component. However, since the electron

component is larger, the emitter current will be dominated by this component.

The base is usually very thin, and thus in steady state, the minority carrier profile for the base is represented by a straight line as shown in Figure 2.2 [SeSm87].



The highest concentration of electrons is at the emitter base junction, and is proportional to e^{V_{be}/V_T} . The term V_T is the thermal voltage, approximately 25mV at room temperature. The reason for the zero concentration at the reverse bias junction is due to the collector voltage which causes the electrons at this end to be swept into the collector.

The electrons injected from the emitter into the base diffuse through the base towards the collector due to the tapered minority concentration profile. The diffusion current is proportional to $nb(0)$, the

electron concentration at the base emitter junction, and inversely proportional to the width of the base.

Some of the electrons injected into the base recombine with the majority carrier holes. However since the base is very thin (typically $1\mu\text{-}2\mu$), recombination is not a dominant effect and most of the electrons are swept into the collector. In the minority carrier profile, this is shown by a small deviation from the straight line producing a concave shape. The slope of the curve at the emitter-base junction (EBJ) is larger than that at the collector-base junction. The difference represents the number of electrons which recombine in the base.

The electrons that reach the collector base junction will be swept into the collector, since the collector is more positive than the base. These collected electrons will constitute the collector current. It is important to note that the magnitude of the collector current is independent of the reverse bias voltage.

2.1.2.2 Current gain in the Transistor

The collector current may be expressed as

$$i_c = I_s e^{V_{be}/V_T} \quad (2.3)$$

where I_s is the saturation current and V_T is the thermal voltage. The reason for the exponential dependence is that the electron diffusion current is proportional to the $n_b(0)$, which in turn is proportional to e^{V_{be}/V_T} . The saturation current I_s is proportional to the area of the

EBJ. Typically I_S is in the range of 10^{-12} to 10^{-15} A and is a function of temperature [Sze87].

As seen above the base current has two components. The first and dominant component is holes injected from the base into the emitter. The second component is the holes that must be supplied to replace the holes in the base that are lost to recombination. The total base current can therefore be expressed as [SeSm87]:

$$i_b = i_c / \beta \quad (2.4)$$

where β is called the common-emitter gain. This gain factor Beta is constant for the particular transistor and can range from 100 to several thousands, and is inversely proportional to the width of the base W . In the following chapter, one of the experiments is to determine the value of beta for the phototransistor.

The emitter current can be seen to be the sum of the base and emitter currents.

$$i_e = i_c + i_b \quad (2.5)$$

$$i_e = \frac{\beta+1}{\beta} i_c \quad (2.6)$$

or

$$i_c = \alpha i_e \quad (2.7)$$

It can be seen that alpha is a constant for a particular transistor, less than but very close to one. This factor is called the common base current gain.

In discussing the emitter current another term which will arise during the discussion of the phototransistor is that of emitter efficiency. Emitter efficiency, γ , is defined as the emitter electron current as a ratio of total emitter current. Clearly from the above discussion γ should be close to one (since emitter current is dominated by electron current).

2.1.2.3 High Frequency Bipolar Model Parameters

In order to discuss the phototransistor operation in later sections, it is useful to provide a model of the bipolar junction transistor. The small signal model for a bipolar junction transistor is given in Figure 2.3 and is called the hybrid pi model [HaNe76].

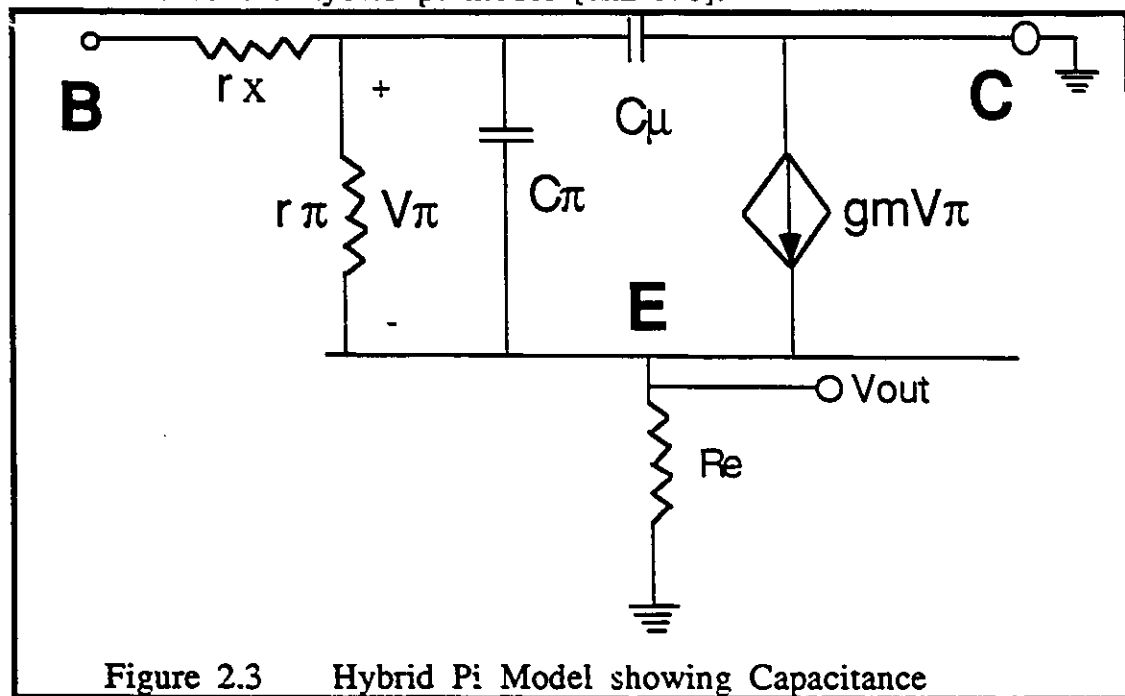


Figure 2.3 Hybrid Pi Model showing Capacitance

The term r_{π} represents the small signal input resistance, between base and emitter and is given by [SeSm87]

$$r_{\pi} = \beta / g_m \quad (2.8)$$

The term g_m , which is called the transconductance is given by :

$$g_m = \frac{I_c}{V_T} \quad (2.9)$$

The equivalent circuit shown applies at a particular bias point, since the two parameters g_m and r_{π} depend on the I_c value. The r_x term models the resistance of the Si material of the base region. Its presence is felt at high frequencies where the input impedance of the BJT becomes highly capacitive.

The more complete hybrid pi models include the two junction capacitances and an output resistance. For a device operating in the active mode (as the photoreceptor is) the forward biased EBJ displays two parallel capacitances: a depletion layer capacitance C_{je} , and a diffusion capacitance C_{de} . The CBJ junction, which is reverse biased, only creates a depletion capacitance, since negligible current flows through a reverse biased diode. The hybrid pi model operating in the active mode (Common Collector Configuration) and containing these capacitors is shown in Figure 2.3. The EBJ capacitances have been added together and called C_{π} . The CBJ capacitance is called C_{μ} .

The capacitor C_{μ} , although usually smaller of the two capacitances, plays an important role in the phototransistor. Due to the large base collector area of the device, C_{μ} is large and plays an important role in

determining the speed of the phototransistor. This will be discussed in more depth when the speed of the photoreceptor is tested.

2.1.2.4 Operation of Photodiode

When a p-n junction is formed in a semiconductor material, a region containing no charge is formed with a high internal electric field across it. If the junction is illuminated by photons with $h\nu$ greater than E_g (the band gap of Si), an electron hole pair is produced. If an electron hole-pair is generated by photon absorption within this region, then the internal field will cause the electron and hole to separate (Figure 2.4) [WiHa83].

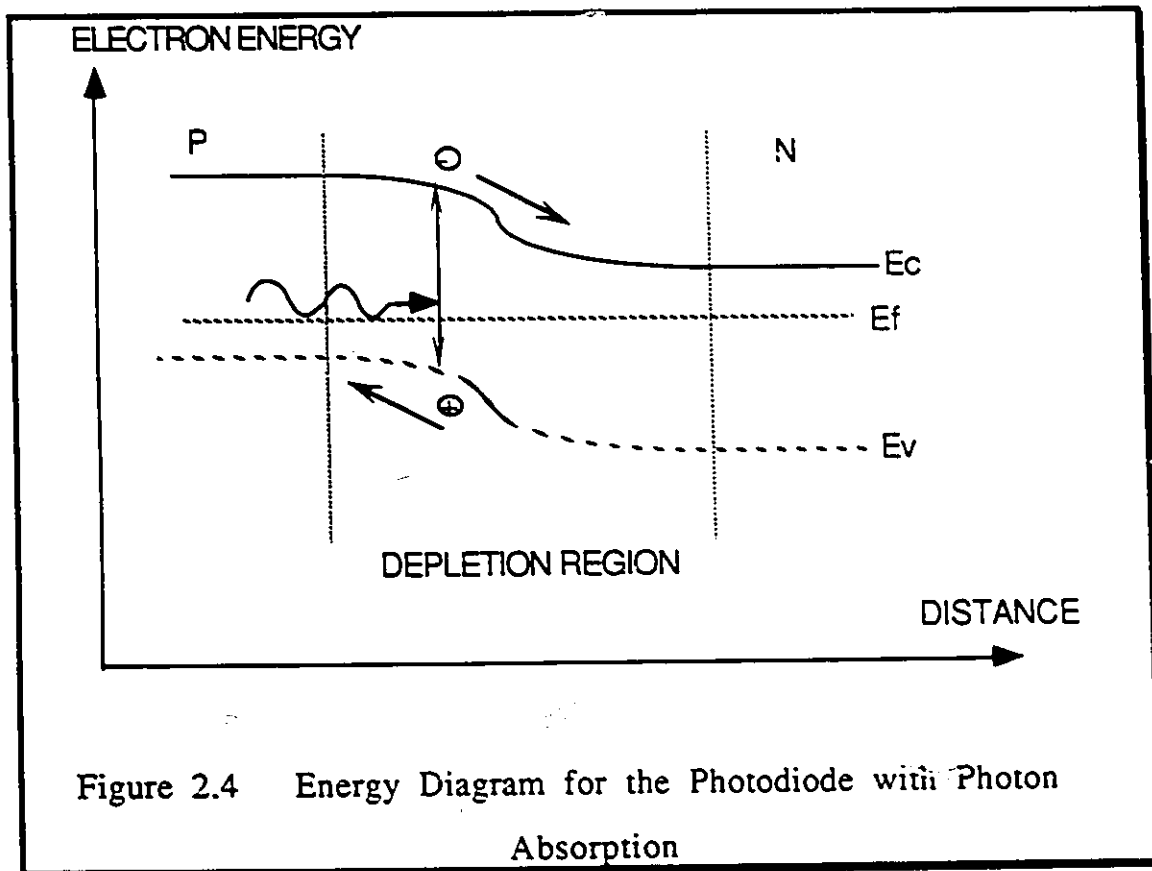
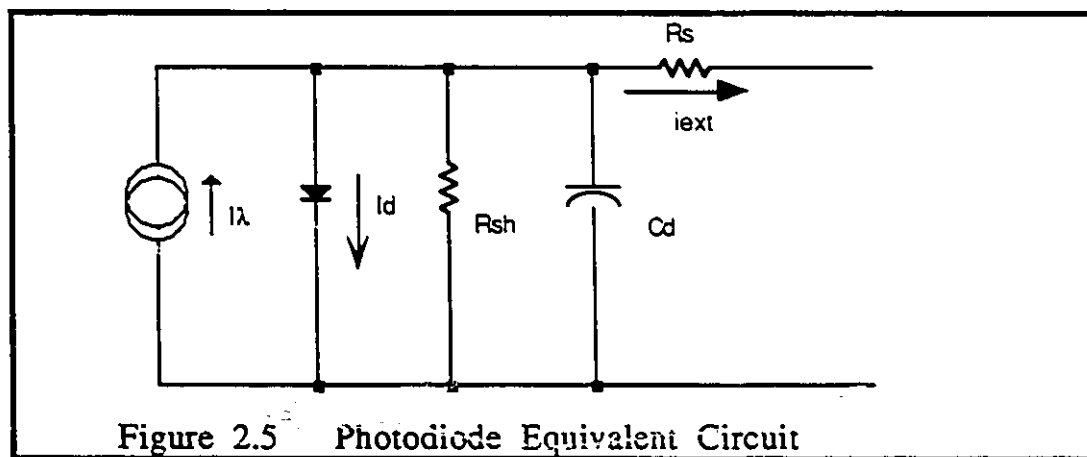


Figure 2.4 Energy Diagram for the Photodiode with Photon Absorption

The junction will also be affected by electron-hole pairs generated away from the depletion region provided they are able to diffuse to the edge of the depletion region before recombination takes place. Only carriers generated within a minority carrier diffusion length of the edge of the depletion region are likely to be able to do this. Still this is a nonegligible contribution.

The photodiode can be represented by a constant current generator (the current flow i_λ being the current generated by light absorption) with an ideal diode across it to simulate the effect of a p-n junction. The internal characteristics of the cell include a shunt resistor R_{sh} , a shunt capacitor C_d , and a series resistor R_s , as shown in Figure 2.5 [Sze81].



If a quantum efficiency, η , for the photon absorption process is assumed, and also that all the incident radiated is absorbed within the cell, then

$$i_\lambda = \frac{\eta I_0 A e \lambda}{h c} \quad (2.10)$$

where I_0 is the intensity of light falling on a cell of area A .

In finding an expression for the photocurrent as a function of light intensity, a low modulation frequency is assumed, so that the effects of the shunt capacitor are neglected. However, this capacitance will become important when dealing with the speed of the phototransistor.

Considering the currents to flow as shown [WiHa83]

$$i_{\lambda} = i_d + i_{sh} + i_{ext} \quad (2.11)$$

also

$$i_d = i_0 \left[e^{\left(\frac{eV_d}{kT}\right)} - 1 \right] \quad (2.12)$$

where i_0 is the diode reverse bias leakage current.

A relatively large reverse bias is usually used to bias the diode. The diode current saturates at i_0 for relatively small values of reverse bias, therefore,

$$i_{\lambda} = i_0 + i_{sh} + i_{ext} \quad (2.13)$$

Now i_0 which is about 10^{-14} A. R_{sh} is several thousand $M\Omega$. i_{sh} given by:

$$i_{sh} = V_d R_{sh} = 10^{-13}A \quad (2.14)$$

so that when it is of the order of a picoamp. it is safe to say that

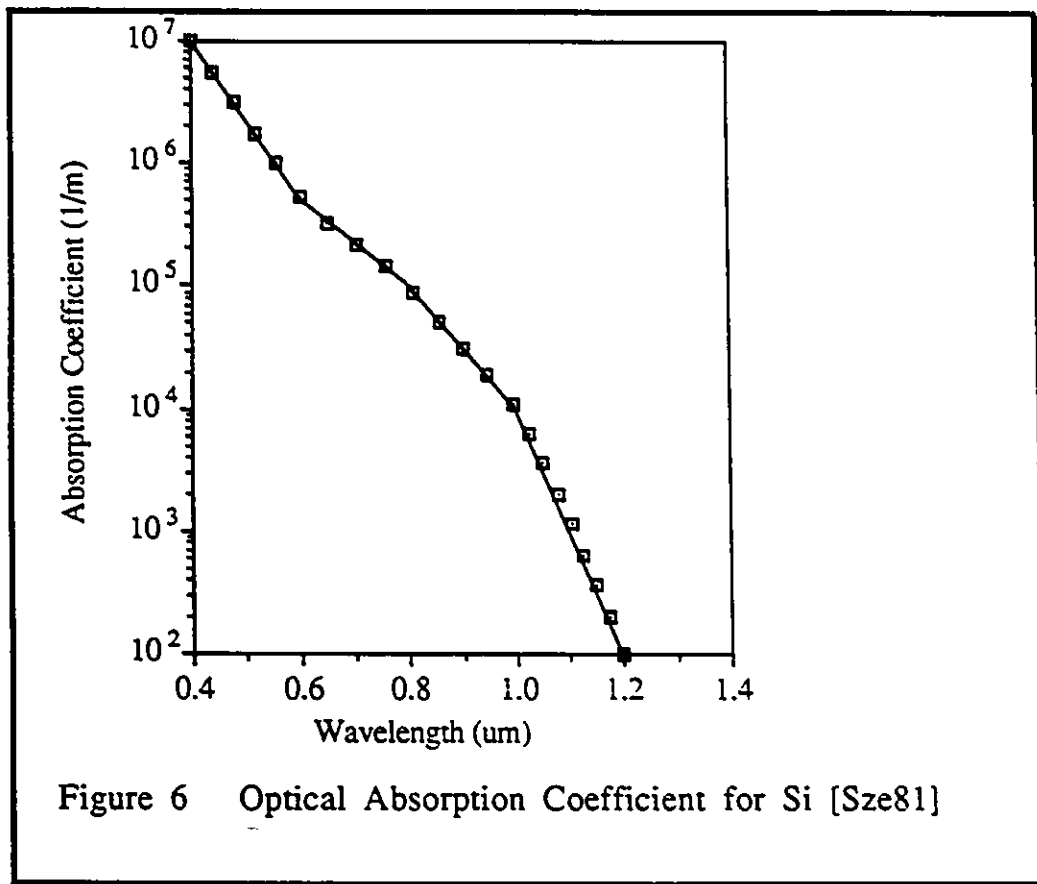
$$i_{ext} = i_{\lambda} \quad (2.15)$$

which produces [WiHa83],

$$i_{ext} = \frac{h I_0 A e \lambda}{hc} \quad (2.16)$$

The important thing to notice is that the current flowing from the photodiode is linearly proportional to the incident light intensity.

This is a good time to discuss one of the important properties of the material used to make the photodiode. The absorption coefficient α (1/m) (a function of wavelength) is a measure of how deep into the material the photon travels before producing an electron hole pair [Stre80]. As shown in Figure 6, for wavelengths longer than the bandgap wavelength (λ), the absorption coefficient is comparatively small. For wavelengths below the bandgap wavelength, α increases rapidly and this implies that most electron hole pairs will be created within a few microns of the semiconductor surface.



For efficient detection the electron hole pairs should be generated either inside or within a diffusion length or so of the depletion region. At short wavelengths, where the absorption coefficient is relatively high (a property of Si), the electron hole pairs will be generated close to the surface. Therefore, in order to have good short wavelength response, the p region should be made as thin as possible. However, at the upper wavelength range of the detector the absorption coefficient is relatively small and a wide depletion region is necessary for good spectral response. This will be discussed further when the phototransistor is tested.

2.1.2.5 Response time of the Photodiode

As discussed above the photodiode is formed between the base and collector of the phototransistor. This large area introduces a large capacitance value which will effect the speed of the device (one of the experiments in the following chapter). Other effects also will limit the performance of the photodiode, thus directly reducing speed of operation of the phototransistor. Some of these limiting factors are discussed here.

Diffusion of carriers is inherently a slow process, the time taken for a charge carrier to diffuse a distance d may be expressed as :

$$t_{\text{diff}} = \frac{d^2}{2D_c} \quad (2.11)$$

where D_c is the minority carrier diffusion coefficient [Sze81].

To ensure that as few as possible carriers are generated outside the depletion region a wavelength related inversely to the absorption coefficient. However, at wavelengths near the bandgap limit, the speed of detection of some of the optically generated carriers will then be limited by the diffusion time. Those carriers generated within the depletion region respond rapidly, while those outside the depletion region give rise to a large fall time response.

In the presence of high electric fields (10^5 V/cm) [Sze81], the drift velocities of carriers tend to saturate. If the electric field is constant, it can be assumed that the carriers move with a constant velocity, V_{sat} . The longest transition time will result when carriers are generated near one edge of the depletion region. In this case, the carrier will have to travel the full distance of the depletion layer width W . This will require a time [Sze81]

$$t_{drift} = \frac{W}{v_{sat}} \quad (2.12)$$

A reverse biased diode exhibits a capacitance caused by the variation in stored charge at the junction. If an external bias voltage is applied on an abrupt pn junction then the capacitance is given by [Sze81]:

$$C_j = A \left[\frac{q \epsilon_s}{2 \left(\frac{1}{N_a} + \frac{1}{N_d} \right) (\phi_i - V_a)} \right]^{0.5} \quad (2.13)$$

where A is the area of the junction (m²)

N_a is the concentration of acceptor ions (m⁻³)

N_d is the concentration of donor ions (m⁻³)

V_a is the applied voltage in V

φ is the zero bias potential V

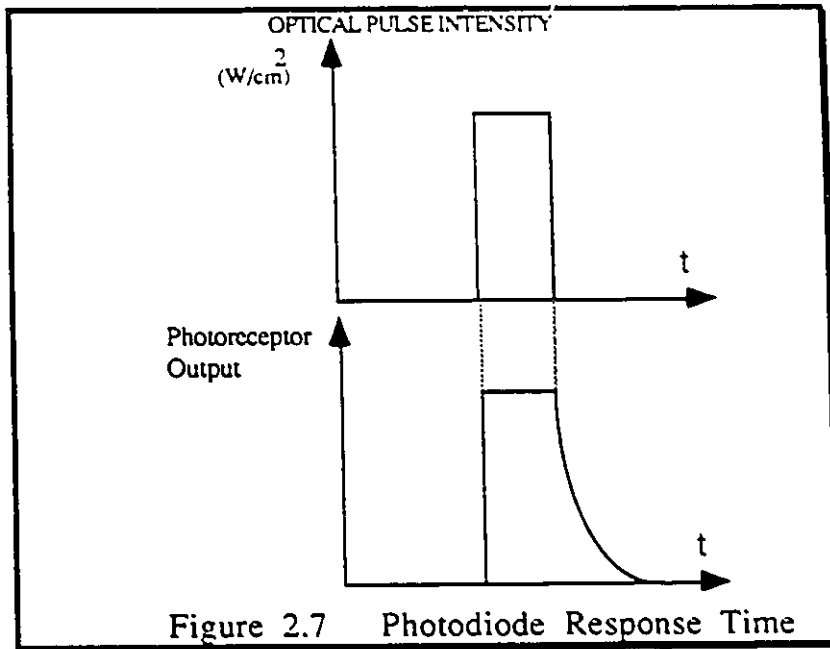
e is the charge of an electron C

ε_s is the permittivity of Si (F/m)

Therefore the output voltage as a function of the frequency f is given by [WiHa89]:

$$V_o(f) = \frac{R_L i_\lambda}{(1 + 4\pi^2 f^2 C_j^2 R_L^2)^{0.5}} \quad (2.14)$$

The bandwidth may be improved by reducing the junction capacitance. However, this would require reducing the diode area A, reducing the doping level N_a, or increasing the reverse bias voltage V_a. However reducing the diode area may make focusing a beam of light on the device a problem. Clearly then the smaller the photoreceptor, the faster the speed: a design tradeoff. An increase in V_a would cause the depletion layer width to increase, and this increases the drift transit time.

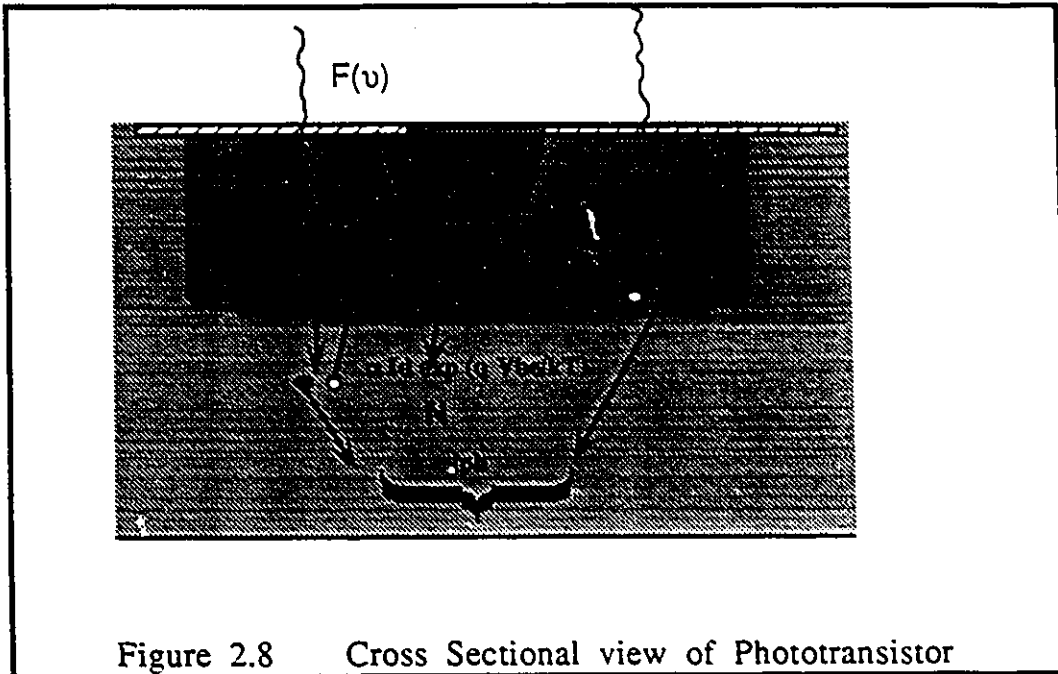


The effect of the junction capacitance as well as the other factor discussed above produces a slow response in the operation of the device when an optical pulse is removed. Experimentally it is expected that this type of response (Figure 2.7 [WiHa83]) will exist when the photoreceptor is tested.

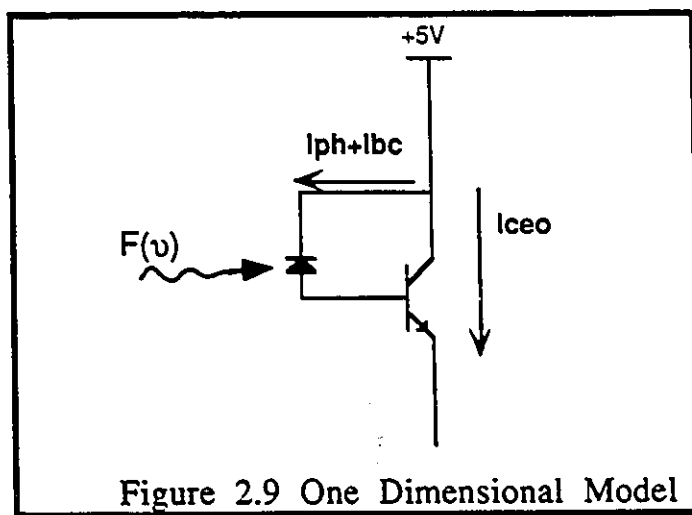
2.1.3 Phototransistor Operation

Illumination of the top surface of the phototransistor (see Figure 2.8) by a uniform photon flux (per unit area) called $F(\nu)$ causes the production of excess electron hole pairs throughout the entire device. Therefore, both junctions EBJ and CBJ act as photodiodes. However, to achieve maximum phototransistor performance, it is desired to suppress the photogenerated leakage current flowing across the emitter-base junction. In practice this is accomplished by making the emitter depth

very thin and relatively small in area (in comparison with the base collector area) [MoChZi71].



Consequently, the phototransistor will be regarded as a photodiode in parallel with the collector base junction of a conventional transistor. This simplistic yet relatively accurate model of the phototransistor is shown in Figure 2.9.



Excess electron-hole pairs are continuously generated on both sides of the collector base photodiode. The excess minority carriers within one diffusion length of the junction move over to the opposite side of the collector base junction giving rise to the photogenerated current I_{ph} . This current is then a majority carrier current and flows in the same direction as the thermally generated leakage current I_{bc} (as shown in Figure 2.9).

The photogenerated current I_{ph} is related to $F(\nu)$ through the quantum efficiency η_{ptd} by the following expression [MoChZi71]:

$$I_{ph} = \eta_{ptd}(\nu) F(\nu) A_b \quad (2.15)$$

where A_b is the area of the base

The incoming radiation $F(\nu)$ then increases the collector base leakage current from its dark value I_{bc} to $(I_{ph} + I_{bc})$. In order to maintain charge neutrality in the base, this majority carrier current must be compensated by injection of minority carriers by the emitter base junction; only then can the net base current equal zero. The current $(I_{ph} + I_{bc})$ is a current source which keeps the emitter-base forward biased. As mentioned in the previous discussions, this keeps the phototransistor in the active mode of operation and gives the device the ability to provide a current gain.

Minority carriers are not only injected from emitter to base but also from base to emitter. The emitter efficiency γ , is used to relate the resulting diffusion currents. I_D represents the reverse saturation current injected from emitter to base. In addition there exists a

leakage component of the emitter current, which is usually represented by means of a diode with a nonideal exponential dependence. The term I_{ET} represents this leakage current. This is the basis for the expression for total emitter current as given by [JoLi68]:

$$I_E = \frac{I_D}{\gamma} \exp(qV_{be}/kt) + I_{ET} \exp(qV_{be}/m'kT) \quad (2.16)$$

The collector current is composed of the current ($I_{ph} + I_{bc}$) plus a fraction α_F of the current injected by the emitter reaching the collector.

$$I_C = \alpha_F I_D \exp(qV_{be}/m'kT) + (I_{ph} + I_{bc}) \quad (2.17)$$

Since the net base current must be zero [MoChZi71],

$$\frac{I_D}{\gamma} \exp(qV_{be}/kT) = \frac{(I_{ph} + I_{bc}) - I_{ET} \exp(qV_{be}/m'kT)}{(1 - \gamma \alpha_F)} \quad (2.18)$$

It is important to note that the forward bias of the emitter junction is controlled by I_{ph} . The output current I_{CEO} (current from collector to emitter with floating base) can be expressed as a function of the photogenerated current [MoChZi71] :

$$I_{CEO} = \frac{(I_{ph} + I_{bc}) - \gamma \alpha_F I_{ET} \exp(qV_{be}/m'kT)}{(1 - \gamma \alpha_F)} \quad (2.19)$$

Since $(I_{ph}+I_{bc})$ represent the base current source an expression $I_{ceo}/(I_{ph}+I_{bc})$ represents the current gain of the phototransistor [MoChi71]:

$$\frac{1}{(1+\beta)} = (1-\gamma \alpha_F) - \frac{\gamma \alpha_F I_{ET} \exp(qV_{be}/m'kT)}{I_{CEO}} \quad (2.20)$$

From the definition of $(\beta+1)$ the output current I_{ceo} is a function of the input photon flux [Sze81]:

$$I_{CEO} = (\beta+1) \eta_{ptd}(v) F(v) A_b \quad (2.21)$$

This relationship shows that the phototransistor has an effective quantum efficiency $(\beta+1)$ times larger than that of the collector base photodiode. It is for this gain requirement that the phototransistor was chosen over its counterpart.

2.1.4 Conclusions

In this section, the theoretical basis for the operation of the phototransistor has been examined. It is clear from equation 2.16 that the device produces an output current which is linearly related to the incoming light intensity. From the model of the phototransistor it can be seen that the device can be represented as a photodiode in parallel with the base emitter junction of a bipolar junction transistor. It was

seen that the current gain of the phototransistor is at least an order of magnitude larger than the base collector photodiode (equation 2.19). From the discussion of the operation of the bipolar junction transistor, a large base to collector was shown to exist and experimentally it is expected that this large capacitance will limit the speed of the device. In section 2.5, the VLSI implementation of the photoreceptor will be discussed. The implementation will be done in Northern Telecoms 3 μ CMOS process.

2.5 VLSI IMPLEMENTATION OF THE PHOTORECEPTOR

2.5.1 Introduction to Fabrication of Photoreceptive Device in CMOS Process

Now that the theoretical background has been discussed, this section will concentrate on the application of these concepts in Northern Telecoms 3 μ CMOS Process. The circuit will be discussed, as well as the VLSI implementation of the phototransistors. Results of simulations will be given, and from this experimental results will be predicted. Finally, practical design considerations implemented will be given .

2.5.2 VLSI Implementation of the Photoreceptor

A 3x3 grid of photoreceptors similar to the ones described by [BeJuMi90] was implemented. It is these photosensitive devices that will be used not only as optical inputs to the photosynapse, but also to program the synaptic weights for large programmable neural networks.

The device consists of a vertical bipolar junction transistor whose output is connected to four diode connected MOSFETs. The current gain on I_{ph} is provided by the BJT operating in the active region. This current flows through a chain of diode connected MOSFETs. These devices (NE in Figure 2.1) provide an output voltage proportional to the logarithm of the current flowing through them.

In the Northern Telecom 3u CMOS process Using Northern Telecoms 3u process the layout of the vertical BJT consists of n+ diffusion in a p-well, the collector is the bulk n type substrate. This transistor is operated in an open base configuration, with the collector tied to VDD. The circuit schematic, and cross sectional layout are given in Figure 2.10 .

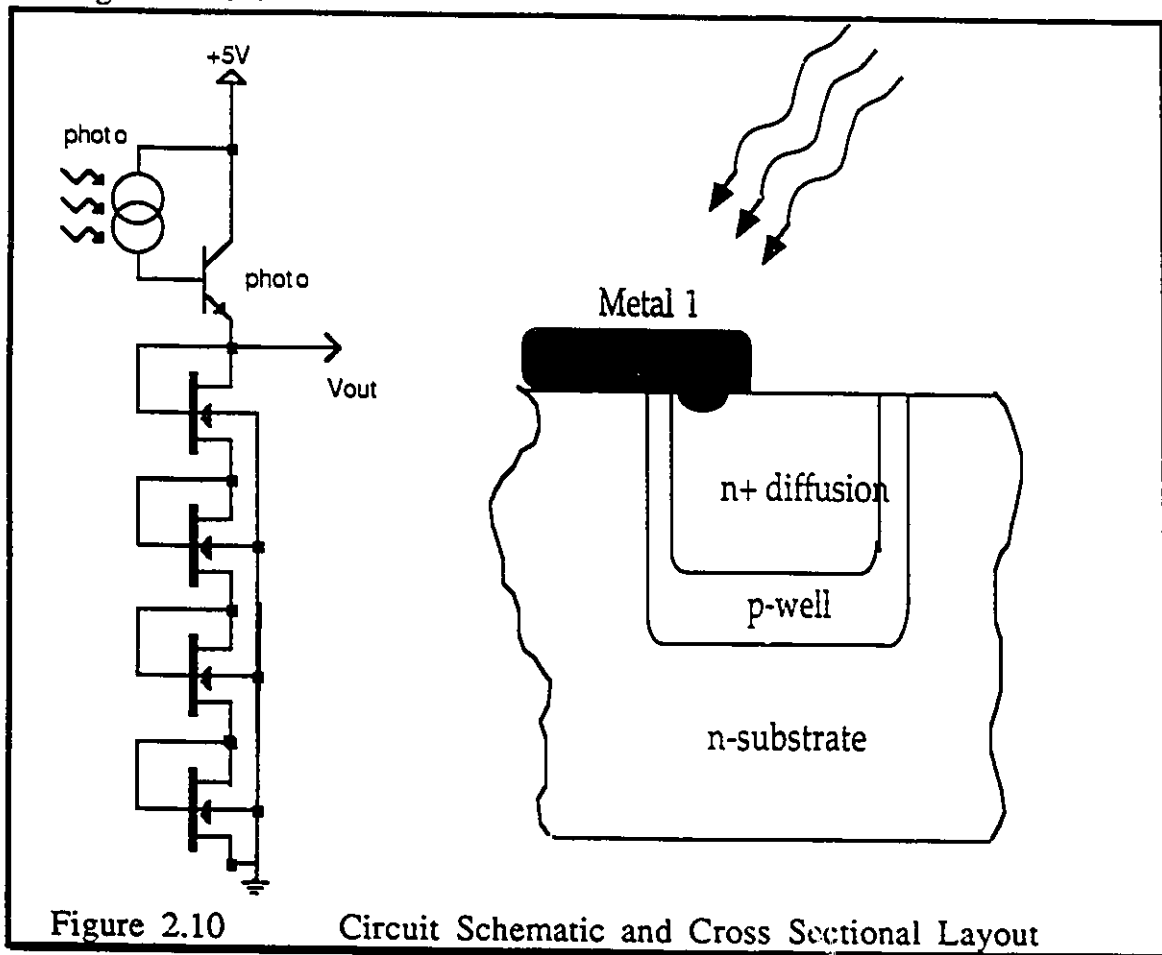


Figure 2.10

Circuit Schematic and Cross Sectional Layout

In this implementation, the emitter is formed by the diffused n+ region, and the base formed by the p-well. In the process used, the substrate is connected to power (+5V), hence biasing the collector node. The diode connected MOSFETs are implemented using minimum dimensions, with their substrates grounded.

From measurement reports [Mead87] it is expected that the current from the BJT will be in the range 10^{-14} to 10^{-10} Amps, for dark to very bright illumination. This current is fed directly to a chain of MOSFETs, biasing them in the sub-threshold region [Mead89]. In this region, the transistors have the following relationship :

$$I_{D_s} = I_0 \frac{W}{L} \left[e^{(\kappa V_g - V_s)/V_t} - e^{-(\kappa V_d - V_s)/V_t} \right] \quad (27)$$

where

$$V_t = \frac{kT}{q}$$

This means the gate source voltage varies with logarithm of drain current(which is linearly proportional to incoming light intensity).

Measurements at other labs [Mead87] with similar devices have shown the photoreceptor was receptive over several orders of magnitude. Simulations using SPICE3 , model level 2 (input file is given in Appendix D) have been carried out and the results are given in Figure 2.11. With a chain of 4 MOSFETs, the photoreceptor functioned over 3 orders of magnitude of light intensity.

The first order value of nV_T for the 4 transistors, is about 2.8 volts. The connection of the substrates to ground (a natural connection in a common p-well) allows the body effect to increase this voltage range over the first order value of nV_T . The output range taking into account the body effect, increases to over 4.5 volts. It is predicted that the device will exhibit this output voltage, when fully illuminated.

The photoreceptors operation depends on photons with energies greater than the band gap of silicon create electron hole pairs as they are absorbed. Holes are collected by the p-type base of the npn photoreceptor, lowering the energy barrier from emitter to base, and increasing the flow of electrons from emitter to collector.

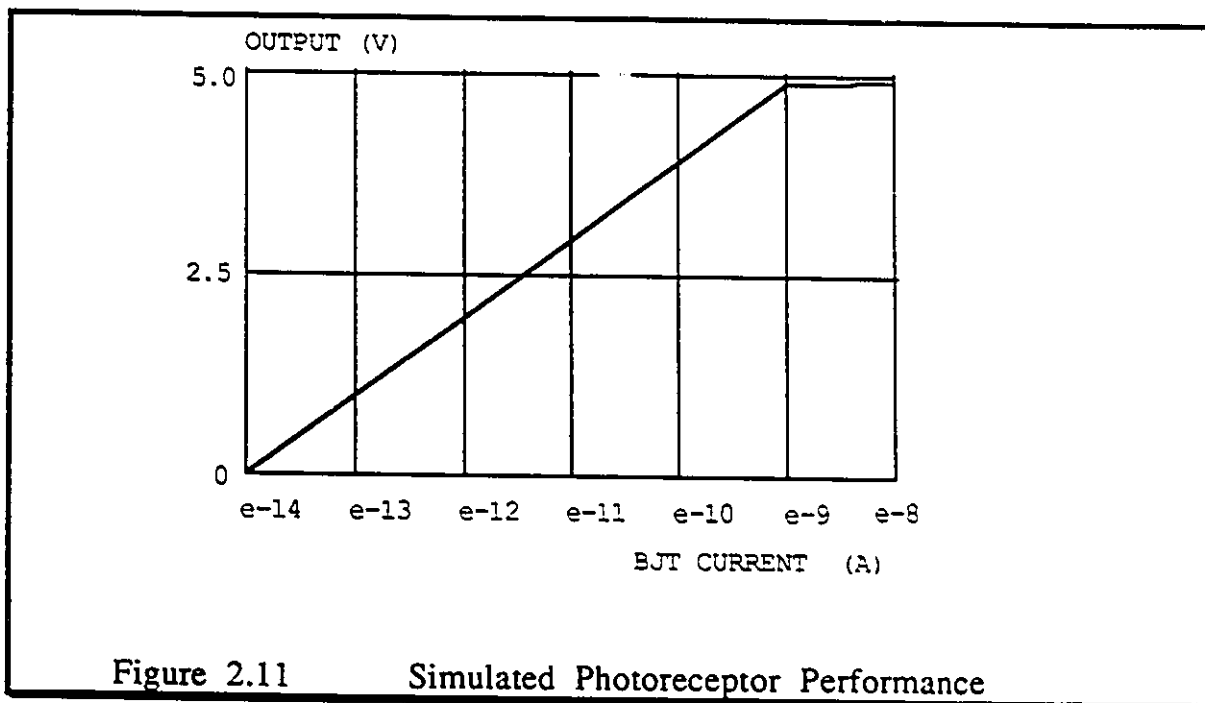


Figure 2.11 Simulated Photoreceptor Performance

The gain of this process is determined by the number of electrons that can cross the base before one electron recombines with a hole in the base. Although the concept is relatively simple to understand,

certain design considerations are implemented to ensure accurate results in the matrix layout of photoreceptors.

A diagram of the photoreceptor which was originally implemented is shown in Appendix A. Clearly, this diagram shows the large base and emitter regions, as well as the four diode connected MOSFETS in a separate p-well. Since the photoreceptor has undergone several generations of "design versions", a picture of the newest version of the photoreceptor is also included in Appendix A. This photoreceptor has several improvements over earlier implementations. These improvements are discussed below.

2.5.3 Practical Design Considerations

The design process of the photoreceptor took several iterations. When starting the design, only some small test circuitry was available. However, it was soon apparent from these semi-functioning cells that several practical design considerations must be implemented to produce a physically robust realization.

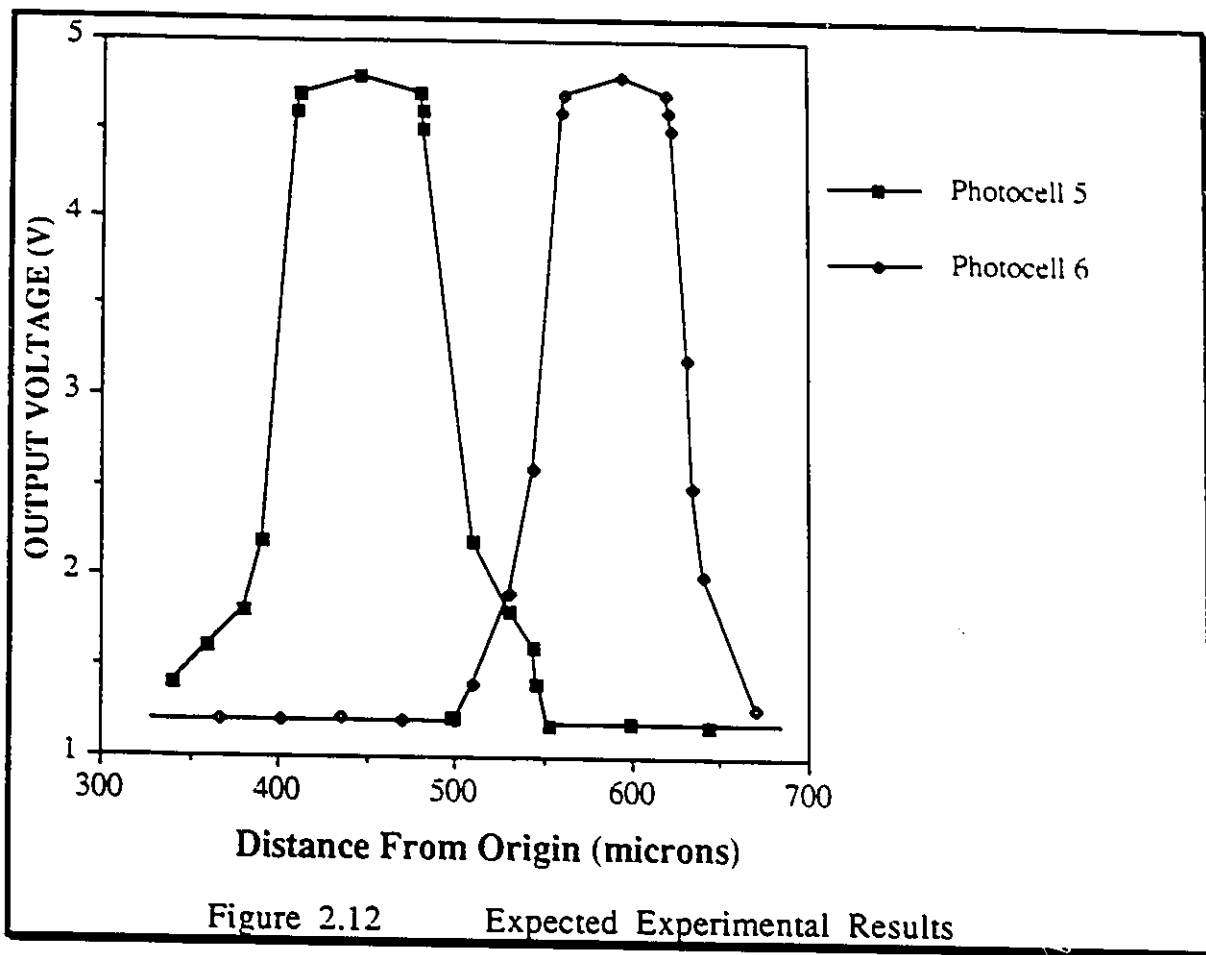
The first consideration is that p-well surrounds each pixel of the photoreceptive grid. The well simply sinks any of these stray carriers (electrons or holes depending on the type of substrate) to ensure that minority carriers from one pixel does not travel into the next pixel, thus producing false inputs. Experimentally this optical crosstalk problem was verified when a photoreceptor cell (with no protection from substrate current) from an initial implementation was tested. By passing a beam of light over two adjacent photoreceptors while monitoring their individual outputs, it can be seen that a great deal of

optical crosstalk exists. The beam of light was of intensity 4.1mW/cm^2 and was powerful enough to saturate the 100μ by 100μ photoreceptor. The beam area was approximately 75% of the area of the photoreceptor (as observed from the eyepiece of the microscope). Figure 2.12 shows the result of such an experiment.

Experimentally we expect the inclusion of p-wells to have a significant reduction in error in the output voltage of the photoreceptors as a light beam is moved across the top row of photoreceptors. The monitored output voltage should clearly show that the implementation of the p-well significantly reduces the errors that are presently occurring from initial attempts at fabrication.

Some of the other practical design considerations that have been implemented are the metal layer which has been placed on the diode connected MOSFETs. The metal is used to reflect the incoming light thus eliminates any erroneous charge carriers created in the transistors due to the exposure to light.

Also, the emitter area has been reduced in order to prevent the EBJ from being influenced from photonical effects. However, since the depth of the emitter region (n-plus) is relatively small, we would not expect the formation of a large number of electron-hole pairs to be formed there. Nonetheless, the reduction in emitter area will further reduce this parasitic effect.



2.3 CONCLUSIONS

The above considerations have been implemented into a device which has been sent for fabrication to Northern Telecom. The device shown in Figure 2.13 has an area of 100u by 100u. Similar devices of differing areas(1,000u² to 100,000u²) have been sent for fabrication. in order to determine if this affects performance.

In the next chapter, the fabricated photoreceptors will be tested. The simulations also predicted a linear relationship between the output voltage and the incident light intensity over four orders of magnitude.

Experimentally we should find a swing of 4.8 volts when saturated with light, as per our simulations. It is also important to ensure that a full swing voltage is produced by the receptor, since it is this voltage which is required for input to the rest of the neural circuitry. We would also expect the photoreceptor to exhibit a long transition time from one illumination intensity to the next. This is caused by the floating base configuration, which accumulates minority carriers with a high lifetime. Successful fabrication and strong experimental results will show the feasibility of implementing these devices in Si as optical inputs to neural networks. It should be noted that a more suitable medium (GaAs) could have been used to implement these devices. However, currently these rules are not available, and hence, a Si process is being used.

III. Characterizing the Photoreceptor

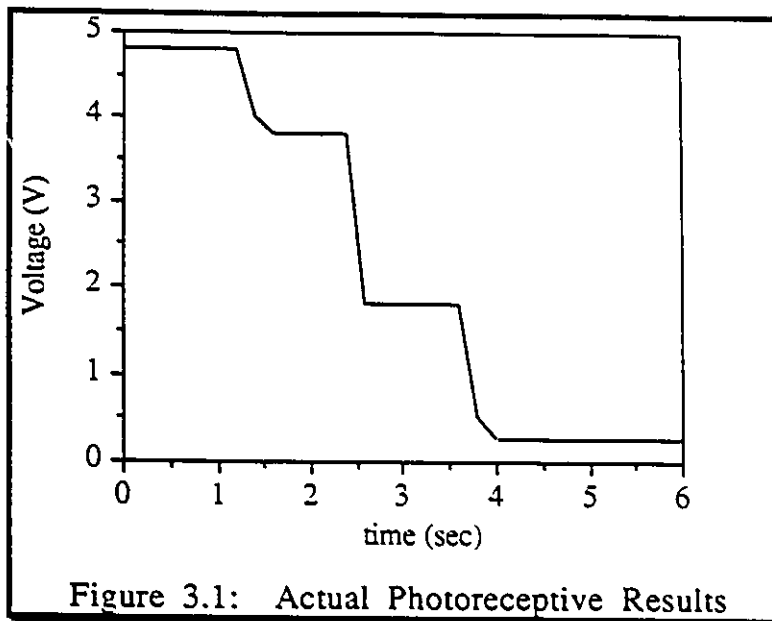
Having received the fabricated integrated circuit from Northern Telecom, the photoreceptor will be characterized by performing a series of experiments. This will enable next generation design changes. The section begins with a description of the initial experimental results to determine the functionality of the device. After this, experiments to determine the current gain factor β , logarithmic dynamic range, and time response will be performed. A discussion of the spectral response is included.

Also included in this section is a discussion of a new photoreceptive configuration that will improve the performance of the device. This configuration makes use of feedback in order to improve the asymmetries encountered in the time response.

3.1 Determination of Output Voltage Range

The purpose of this experiment is to determine the response of the photoreceptive device when exposed to various intensities of illumination.

The test fixture consisted of a Halogen Tungsten light source focused through a Bausch and Lomb microscope (model # OEM). Using the microscope's lens system, the light is focused onto the surface of the integrated circuit (i.c.). The chip is placed in a 40 pin zero insertion force (zif) socket, which is attached to a x,y,z table (Wentworth Laboratories MP900 Probing Station). The x, y, and z positions are locked in order to prevent movement. The power,



ground, and output voltage are connected to the Hewlett Packard Parameter Analyzer (model # HP4145HB).

Spice simulations predict the output voltage will swing from 0V to approximately 5V,

when the incident light is off to very bright (fully saturated) as shown in Figure 2.11. The three discrete values of light intensities (produced by the coherent light source) should be enough to determine the functionality of the device.

The integrated circuit was placed in the test fixture and the Parameter Analyzer was programmed to measure the output voltage as a function of time. The first light intensity was 4.1 mW/cm^2 and the output voltage was measured to be approximately 4.8V. This procedure was repeated for light intensities of 3.0 mW/cm^2 and 1.5 mW/cm^2 respectively. These measurements were performed using a Metrologic Radiometer. Exposure of the chip to no illumination produced an output voltage of 0.4V. These results are shown in Figure 3.1.

The results show that the functionality of the device works as predicted by the simulation. However, there are some interesting observations in the result. Firstly a relatively long time seems to be required for the output to reach a new output voltage, after the illumination was changed. This is due to the delay time required for the Tungsten filament of the light source to change to a new illumination level. Another interesting observation is the output voltage with no illumination incident on the surface of the integrated circuit. This was repeated after the room's ambient light is removed (as best as possible). This "dark voltage" can be explained by the presence of ambient light as well as thermally generated current produced in the device.

In conclusion this experiment has shown that the response of the receptor swings from a dark voltage of approximately 0.4 volts (dark voltage) to 4.8 volts when fully saturated. This is approximately the voltage range predicted by earlier simulations.

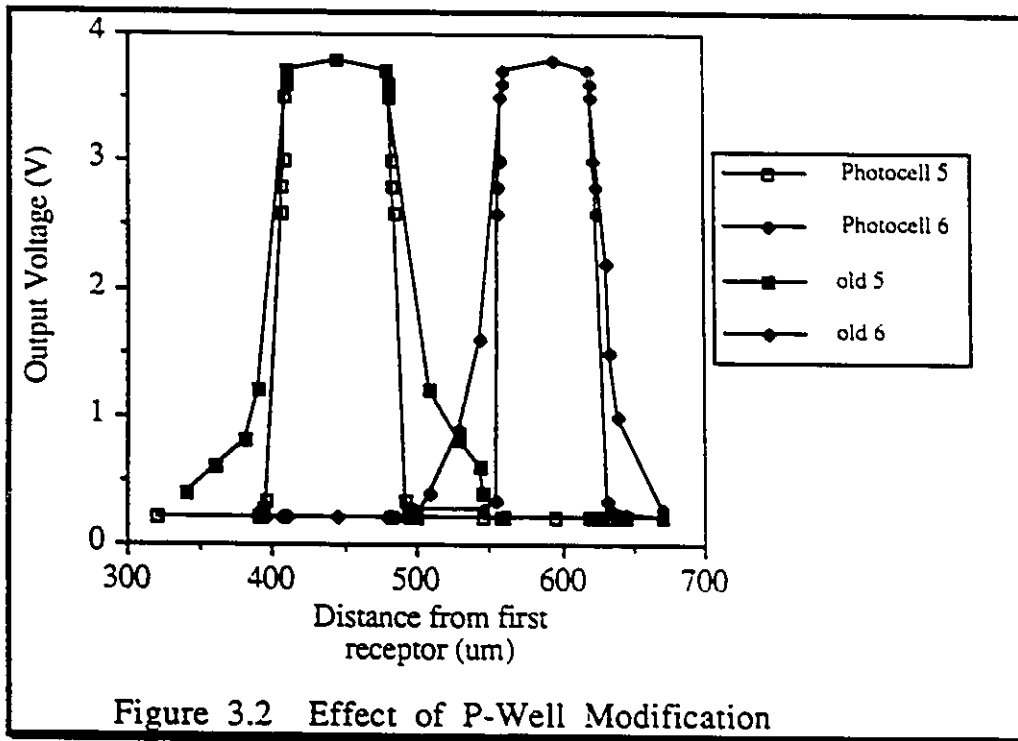
3.2 Removal of Optical Crosstalk by P-Well

The purpose of this experiment is to determine if optical crosstalk between adjacent photoreceptive devices is reduced by the addition of p-wells around individual receptors.

The test fixture consisted of a coherent light source from a Meilles Girot He-Ne laser focused through a Bauch and Lomb microscope (model # OEM). The integrated circuit is placed into a 40 pin zif socket. This socket is attached to an x,y,z table. The table is allowed to move in the x direction ;the y. and z movements are

restricted. The power, ground, and output voltages of two adjacent photoreceptors are monitored by the Parameter Analyzer (model # HP4145B).

The light source was focused on the first photoreceptive element. The incident light intensity was 3mW/cm^2 . The light source was moved in the x direction from one photoreceptive element to the next. The output voltage of both photoreceptors (with p-well protection) is monitored. This is compared to a similar experiment run on photoreceptors with no p-well protection. The output voltage for both is shown in Figure 3.2.



The results clearly show a decrease in the optical crosstalk with the addition of p-well. Since the photoreceptive elements will be used as optical inputs to an integrated circuit, the validity of the

data is crucial. The illumination of one device in a grid (as the photoreceptors are usually arranged) should not induce an output signal from an adjacent device. The addition of a p-well clearly prevents this.

3.3 Characterizing the Vertical BJT

The purpose of this experiment is two part. The first is to extract characteristic I_C vs V_{CE} curves for the npn vertical BJT. This will enable the Spice model card to be modified in order to more closely reflect fabricated values. The second part of this experiment is to determine the variation of β (Section 2.2.2.2) as a function of I_C . Both parts of the experiment were performed on vertical BJT npn transistors on CMCs test strip. These test devices are fabricated similar to the device shown in Figure 2.10.

A H.P. Parameter Analyzer (model # HP4145B) was used to provide voltages and currents to the chip, as well as measure the output current. Since the test strip is not bounded to output pins, probe pads were used to probe into the device. The test chip was placed in a 40 pin zif socket in order to prevent its movement.

The emitter node was grounded. The collector emitter voltage was varied from 0 to 5V, while I_B was varied from 1 μ A to 5 μ A. The collector current I_C was monitored. After several attempts, a set of characteristic curves of I_C vs V_{CE} were obtained, as shown in Figure

3.3. The data was then used to change the model card of the npn transistor. The final Spice deck is given in Appendix III.

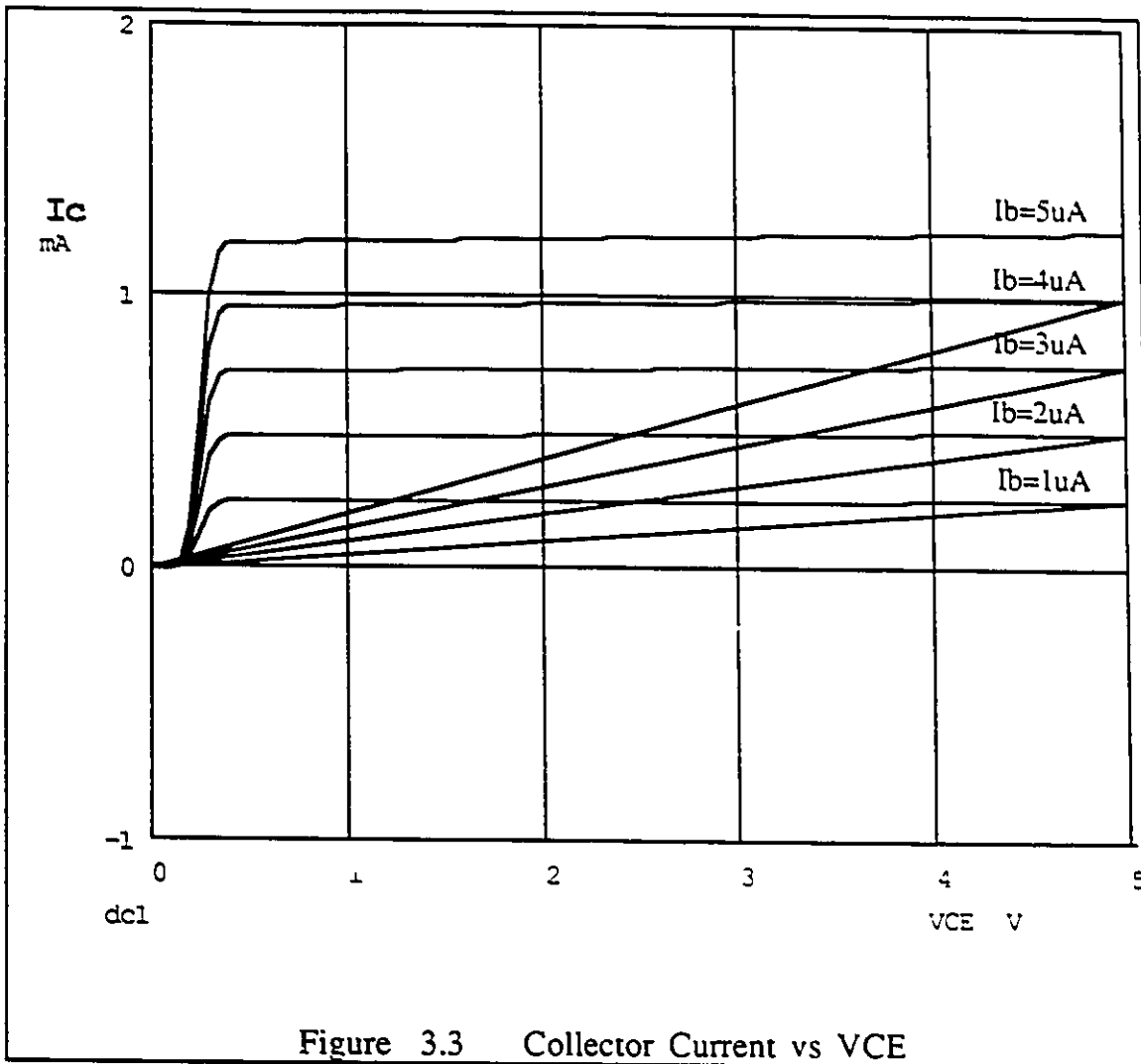


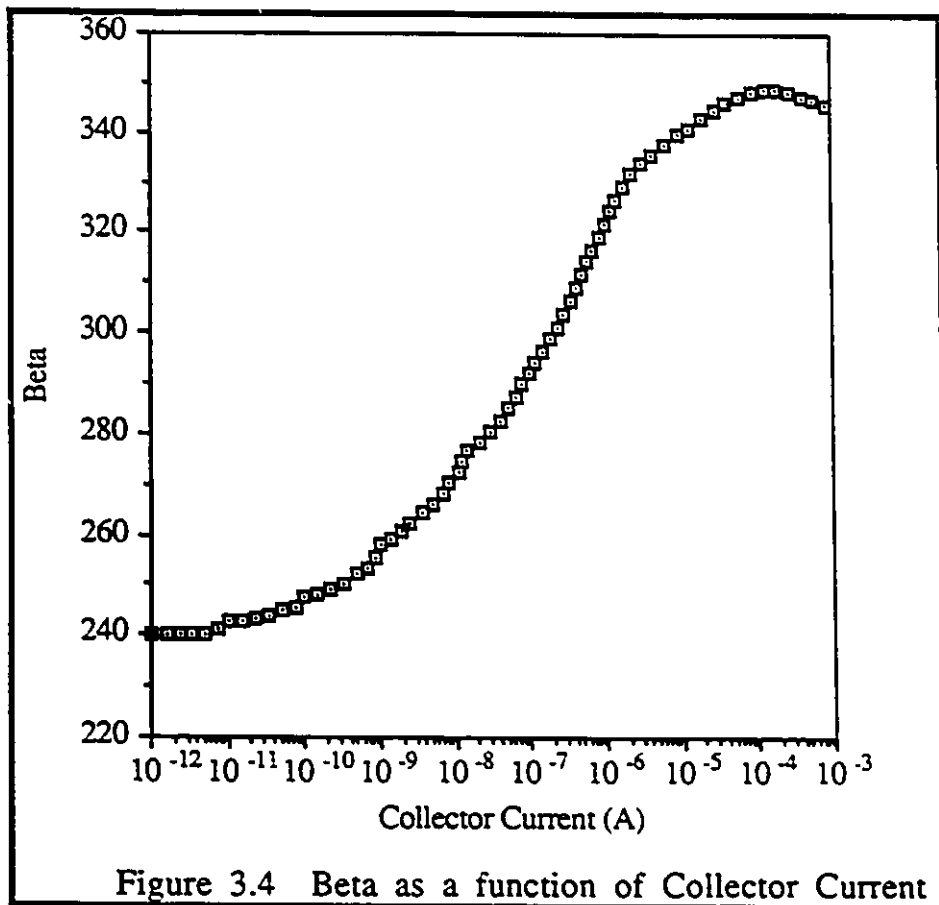
Figure 3.3 Collector Current vs VCE

The purpose of the second part of the experiment was to plot β as a function of I_C . Using a similar test fixture as before, the values of I_B and I_C were monitored while V_{CE} was held constant at 5.0V. The Parameter Analyzer was programmed to calculate the β value for each reading of current. Figure 3.4 shows the obtained results.

For subthreshold currents, the measured β value is approximately 250. As I_C increases, the value of β steadily increases and then for currents above 0.1mA β starts to decrease. Two different effects account for these regions [AnMa88].

At lower current densities, the average minority carrier does not travel as far before recombination occurs. The average minority carrier lifetime decreases. Hence, the recombination in the base is higher. This causes a larger number of holes to be injected into the base from the external circuit (a higher I_b). Thus, the current gain $\beta = I_C/I_b$ decreases with decreasing current levels.

As the current level is increased, a second effect takes over



which causes β to decrease with increasing current. The emitter current consists of holes injected from base to emitter. This hole component which was very small before (and neglected) becomes an increasingly significant component of that current. As a result I_b increases at a higher rate, causing the current gain β to decrease.

In summary the I_c vs V_{ce} characteristics of an actual fabricated transistor were measured. From these curves, the Spice deck parameters were varied in order to compare with experimentally observed values. The current gain β of the transistor was found to be around 250 for "typical" photocurrents. It was also found that β increases with emitter current levels until at very high levels ($>0.1\text{mA}$) the gain rapidly dropped off. The variations for this behaviour were traced back to the lifetime minority carriers in the base (for lower currents) and the significant components of emitter current injected from the base (for higher currents).

3.4 Logarithmic Response of the Photoreceptor

The purpose of this experiment is to investigate the dynamic range of the logarithmic photoreceptor.

Theory (equation 2.16) predicts a photocurrent that is linear with light intensity. In sub-threshold a diode connected transistor gives an output voltage that is linearly related of the logarithm of the current. The tested cell is therefore expected to have a **logarithmic voltage response as a function of light intensity.**

3.4.1 STATIC CHARACTERISTICS

Two diode connected transistors are connected as shown in Figure 3.5a and a photocurrent I_{ph} flows. Calling the output voltage V_{out} , and the voltage in between the two diode connected transistors V_1 , the current can be shown to be

$$I_{photo} = I_o e^{\kappa V_1} = I_o e^{(\kappa V_{out} - V_1)} \quad (31)$$

where the voltages are measured in units of kT/q . Eliminating V_1 between these two equations, we obtain [Mead90]:

$$V_{out} = \frac{\kappa+1}{\kappa^2} \ln \frac{I}{I_o} \quad (3.2)$$

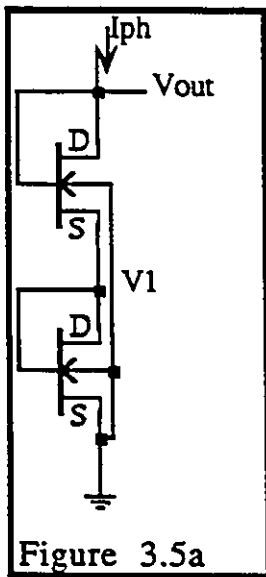


Figure 3.5a

Therefore, the output voltage of the receptor is expected to follow a logarithmic relationship to the incident light intensity. It is the range of this relationship that will be experimentally verified.

3.4.2 EXPERIMENTAL PROCEDURE

The test fixture consisted of a Halogen Tungsten light source focused through a Bausch and Lomb microscope (model # OEM). Using the microscope's lens system, the light is focused onto the surface of the integrated circuit (i.c.). The incident light intensity

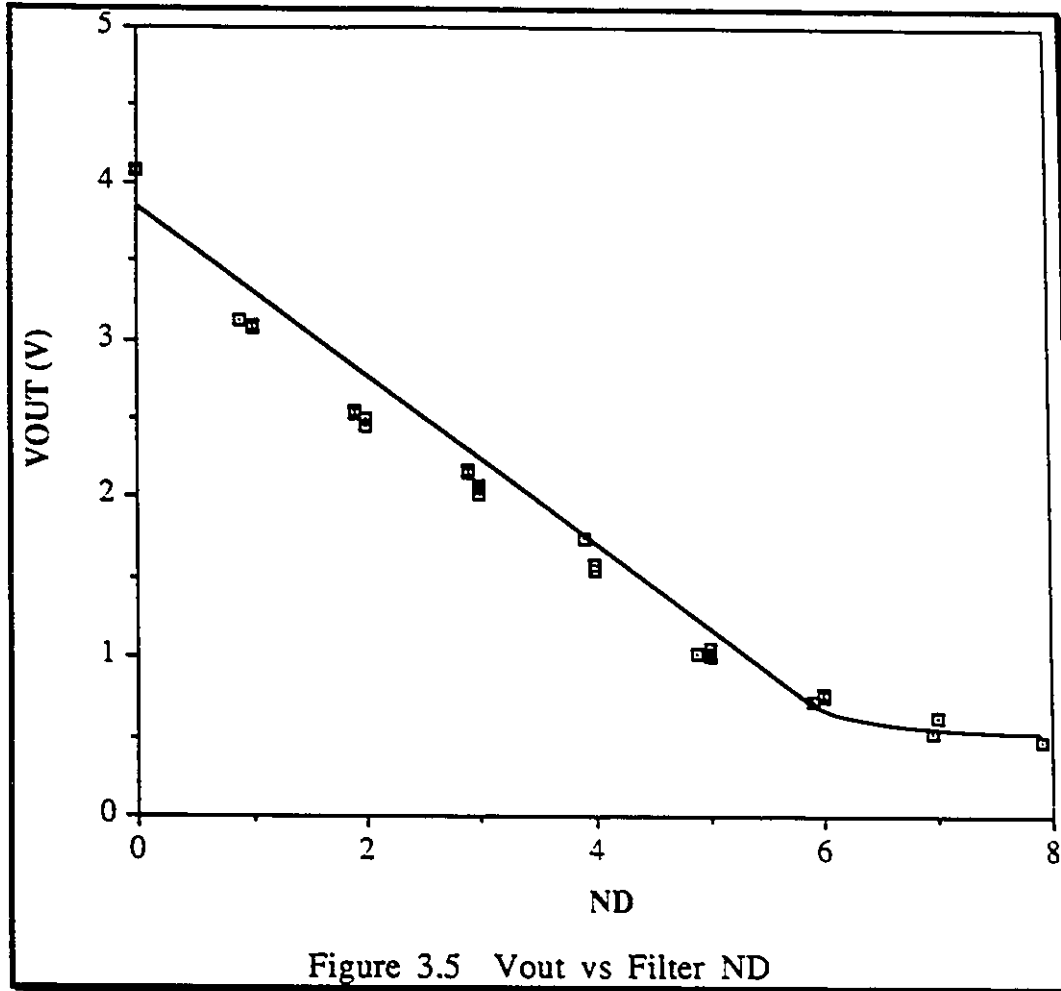
is measured using a Metrologic Radiometer to be 3mW/cm^2 . The chip is placed in a 40 pin zif socket, which is attached to a x,y,z table (Wentworth Laboratories). The x, y, and z movements are locked in order to prevent the distance from the microscope to the surface of the integrated circuit from varying. The power, ground, and output voltage are connected to the Hewlett Packard Parameter Analyzer (model # HP4145B). A variety of neutral density filters were used to cut down the light intensity(ideally) by factors of 10^{-ND} , where ND is the neutral density factor of the filter. The filters were of ND's = 0.9, 1.0, and 2.0. Combinations of these filters could be used to get data between ND=0 and ND=7.9 for the experiment (2 of ND=2, 3 of ND=1 and 1 of ND=0.9) [Mead90].

The filters were placed between the light source and the surface of the integrated surface. The output voltage was monitored as combinations of ND filters from 0 to 7.9 were used. The experiment was repeated several times. These readings are plotted in Figure 3.5b .

There were sources of error introduced by the calibration of the filters (calibrated to 5% by the manufacturer).

3.4.3 EXPERIMENTAL RESULTS

The dynamic range of the logarithmic photoreceptor extended to about seven orders of magnitude. The response was linear over approximately five orders of magnitude and had a gradient of 0.5V per decade as shown in Figure 3.5.



It can also be seen that the curve flattens out at 0.4 volts, this suggests that there is a leakage current flowing in the phototransistor in the dark.

3.4.4 CONCLUSION

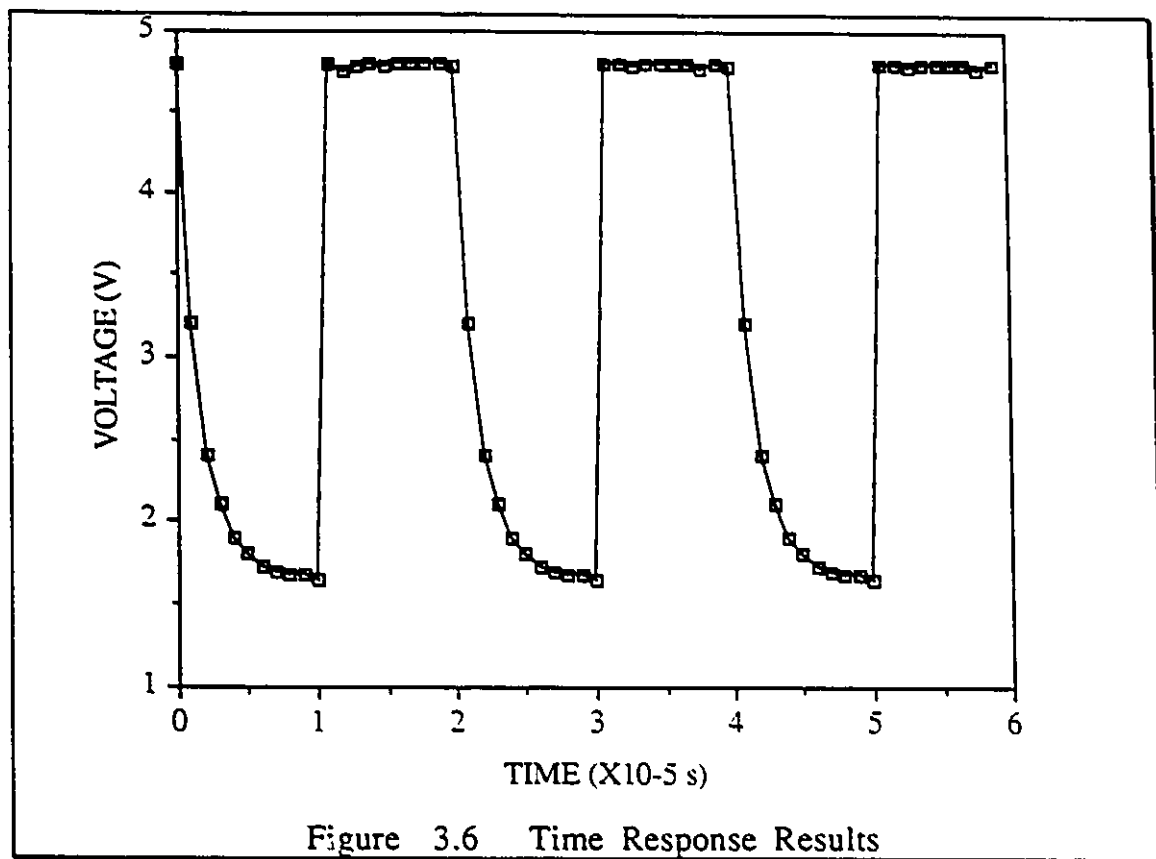
The dynamic range of the logarithmic photoreceptor extended to about seven orders of magnitude. The response was linear over approximately five orders of magnitude and had a gradient of 0.5 v per decade. Since the circuit had such a large dynamic range, it shows great promise as a very sensitive and wide range light detector for VLSI circuitry.

3.5 Transient Response

The purpose of this experiment is to measure the transient response of the photoreceptor. Any asymmetries between the rise and fall times will be identified and explained.

The equipment used in this experiment is as follows. The external input to a Melles Griot He-Ne laser was fed a 1V, 20us square wave through a simple filtering circuit as shown in Appendix IV. The output light intensity was measured to be 4.1mW/cm^2 . The integrated circuit was placed in a 40pin zif socket. The HP Parameter analyzer was used to monitor the output voltage of the photoreceptor, as well as provide power and ground voltages to the chip.

The strobe action was started, and the output voltage was monitored and plotted in Figure 3.6. The figure clearly shows large asymmetries between the rise and fall time. A sudden increase in the illumination resulted in a sudden increase (negligible delay) in the output voltage of the device. The turn on time of the transistor is determined by the rate at which charge is dumped into the base which is directly related to the light intensity.

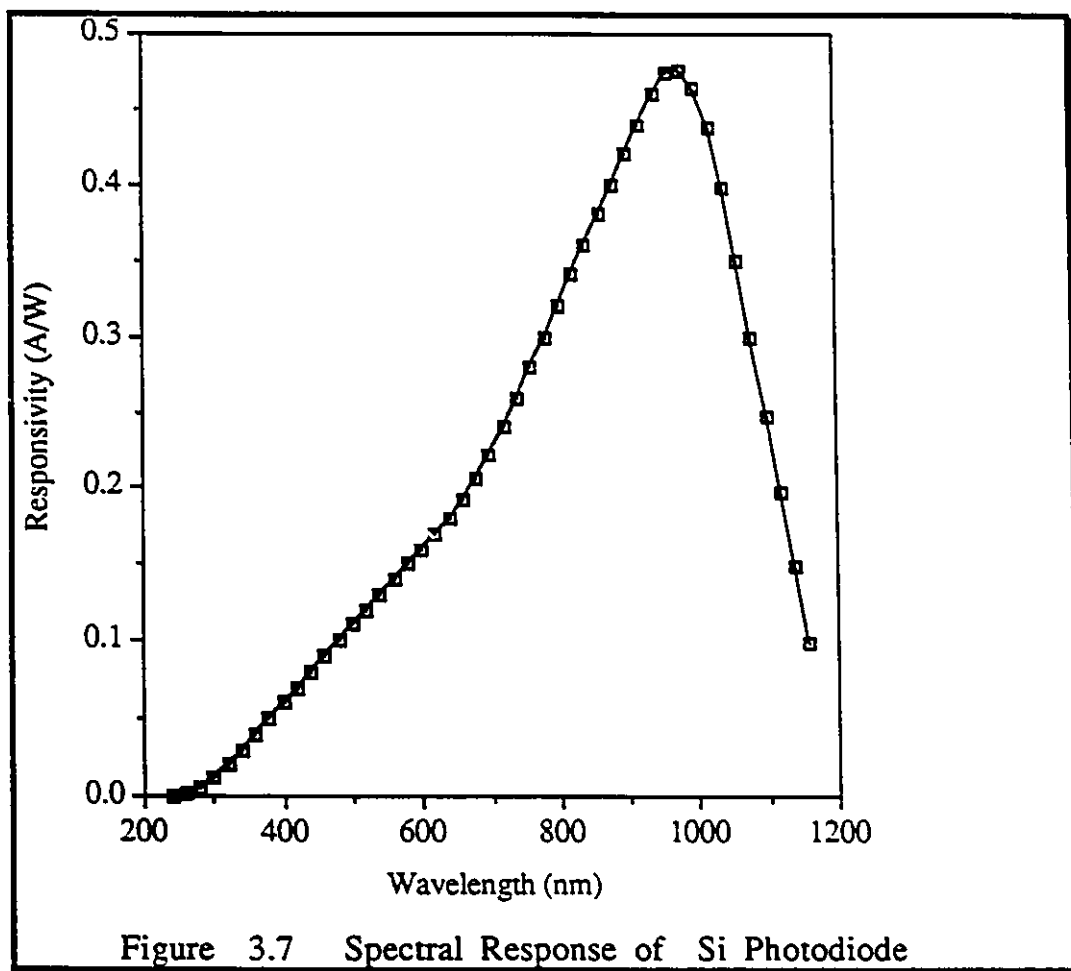


However, the response to cessation of light followed an exponential decay in the order of 10 μ s. This type of response in Section 2.2.2.5 where the response time of the photodiode was discussed. It was shown that the discharge of the large capacitance (associated with the large p-n junction area) produces a large fall time.

In summary, it has been shown that an asymmetry is present between the rise and fall times in the response of the photoreceptor. The slower off time of the transistor is caused by the large collector-base capacitance.

3.6 Spectral Response

A typical spectral response [Sze87] as a function of the light frequency is shown in Figure 3.7. The responsivity (y-axis) is the ratio of induced photocurrent (A) to the optical power used measured in W. The most important feature of the graph is the peak near the $\lambda=900\text{nm}$.



The spectral response is a result of the position of the electron-hole pairs produced in the transistor. This was discussed previously.

when the absorption coefficient α ($1/m$) was shown to be a property of the material and related to the frequency of the incident light. This is discussed in more detail in Chapter 2.

When electron-hole pairs of high frequency are produced, they are located just inside the incident surface of the transistor. The energy barrier is not lowered by the electron-hole pairs and as a result the electrons are not allowed to move into the base. The electron-hole pairs are therefore not effective at increasing the collector current. More of the electron-hole pairs are created when the frequency is lowered which causes the light's penetration depth to increase. As this occurs the collector current rises, thus producing a higher output voltage. When the frequency of the light is lowered further the light has insufficient energy to produce enough electron-hole pairs to lower the base potential, and the collector current is again small, reducing the output voltage.

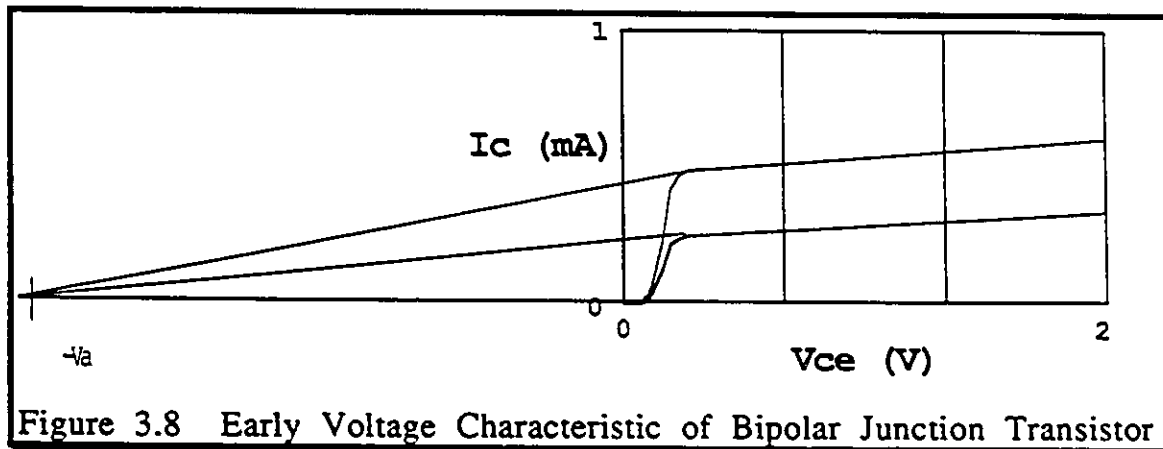
From the results presented by [Sze81], it is clear that the spectral response is not a factor to be ignored in implementing this device in different applications. From the graph it can be seen that the measured peak in the frequency response or best operating range is between 900nm and 950nm in wavelength.

3.7 DETERMINATION OF THE EARLY VOLTAGE

In this section, the cause of the Early voltage will be discussed, as well as the experimentally obtained values of the Early voltage. Future sections value will show how the Early voltage of the phototransistor limits the performance of the photoreceptor.

3.7.1 Experimental Setup

The I_C vs V_{CE} characteristic curves (previously obtained), though still straight lines, have finite slopes. In fact, when they are extrapolated, the characteristic lines meet at a point on the negative V_{CE} axis (Figure 3.8), at $V_{CE} = -V_A$. The voltage V_A , a positive value, is a parameter for the particular BJT, and is called the Early Voltage. In order to determine the value of this voltage, experimental curves were compared with spice simulations. The spice deck parameters were changed in order to obtain the best correlation. The vertical BJT used produced a $V_A=100V$. The experiment is



repeated several times since the measurement of the current must be very accurate: the current curve is very flat and hence the projected point is going to be very sensitive to small errors in the measured current.

3.7.2 Cause of the Early Voltage

If at a given value of V_{be} , increasing V_{ce} results in the increase of the width of the depletion region of the junction. This results in a decrease in the effective base width W as shown in Figure 3.9 [SeSm87]. This is the Early Effect. Since I_s is inversely proportional to the base width, I_s and I_c will increase proportionally.

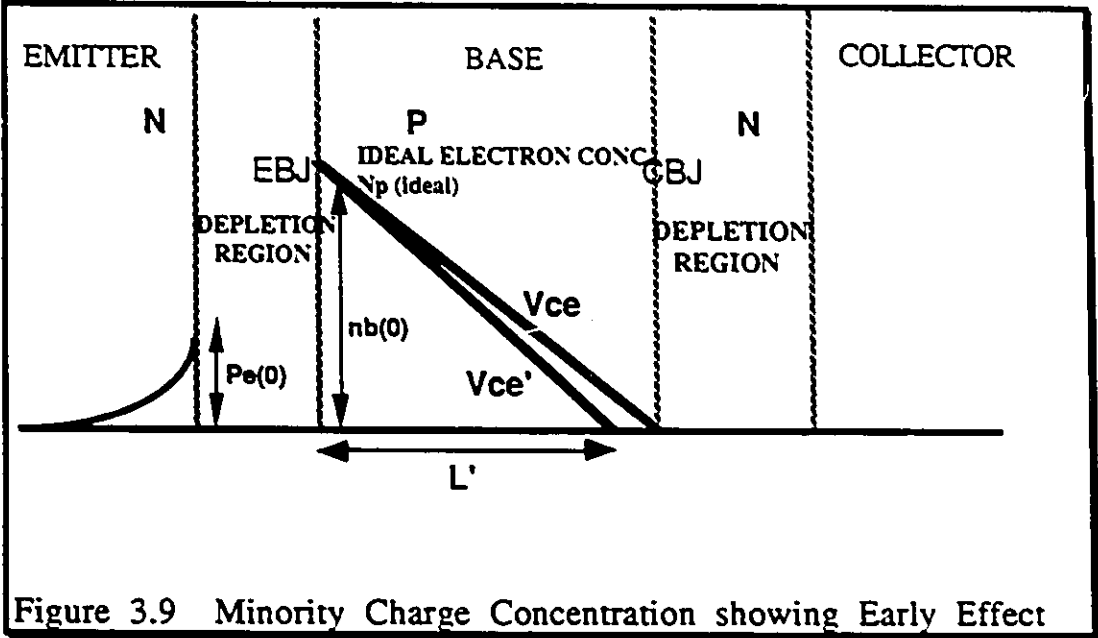


Figure 3.9 Minority Charge Concentration showing Early Effect

The linear dependence of I_c on V_{ce} can be accounted for by assuming that I_s remains constant and including the factor $(1 + V_{ce}/V_A)$ in the equation for I_c :

$$i_c = I_s e^{(V_{bc}/V_T)} \left(1 + \frac{V_{ce}}{V_A} \right) \quad (3.3)$$

3.7.3 Effects of the Early Voltage

The reason that this section was included was to point out one of the weaknesses of the standard configuration of the photoreceptor. The Early effect in the phototransistor will cause an additional increase in the collector current as the V_{ce} is increased. However, an increase in the output voltage (due to illumination by a more powerful light source) is in fact an increase in V_{ce} . Clearly then the relationship between incident light intensity and collector current is not linear as predicted by equation 2.16. For this reason, a new VLSI realization which deals with minimizing this effect is presented in the following sections.

3.8 Various Photoreceptor Configurations

In this section, a photoreceptor configuration designed to improve the performance of the photoreceptive element is described. A discussion of the circuit operation as well as simulated results are presented.

3.8.1 Active Feedback Photoreceptor

Before discussing the design of this new photoreceptor configuration, the author would like to give some background on how the ideas used in this section came about. During a visit of Meads lab at Caltech in April 1990, the author had the opportunity to discuss how the VLSI Research Group at the University of Windsor was using the photoreceptor. During this discussion, we discussed improvements that could be made to the design of the device. At the time they were investigating using "some" sort of feedback to reduce the effect the Early voltage had on the photoreceptor. Since that visit the author have been investigating this approach at our labs.

The logarithmic photoreceptor described in Chapter 2 and 3 makes use of a bipolar phototransistor in order to produce a voltage which is logarithmically related to the incident light intensity.

One disadvantage of the standard logarithmic photoreceptor is that the presence of Early effect causes a variation in the light intensity/ collector current relationship, when the output voltage (V_{ce}) is changed. One possible solution to this problem is to

incorporate active feedback as shown in Figure 3.10 [Mead90]. Now the collector-emitter voltage of the phototransistor is held constant. This is due to the fact that V_1 (V_{ce}) is connected to the gate of a transistor which is forced to supply a constant amount of current as determined by V_b .

3.8.1.2 Static Response Properties

The gain of the standard output (V_2) of this circuit is simply $kT/\kappa q$, since the equation for Q_1 is

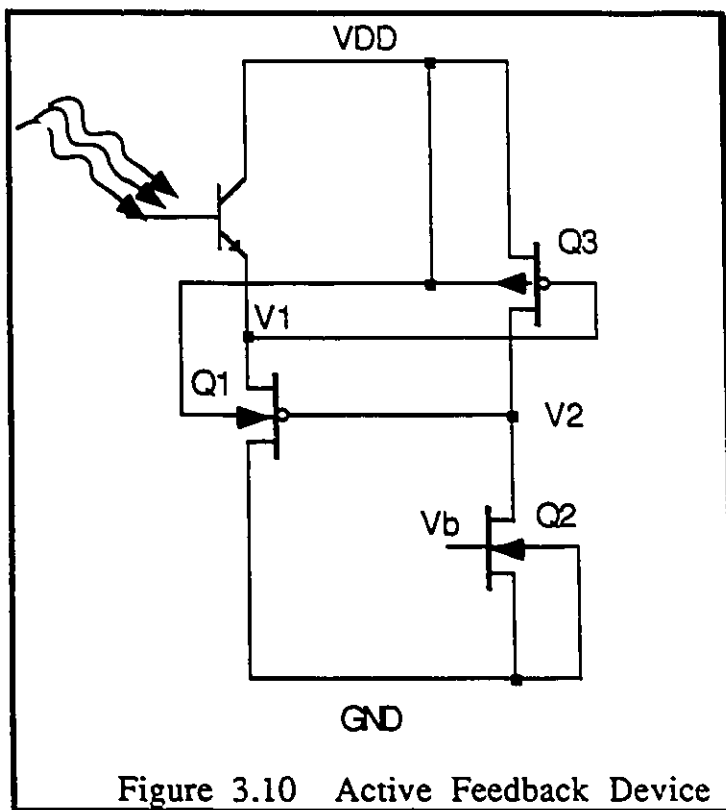
$$I = I_o e^{(\kappa V_2 - V_1)/V_o} \quad (3.4)$$

where V_o is kT/q and therefore,

$$V_2 = \frac{kT}{q\kappa} \ln \left(\frac{I}{I_o} \right) + \frac{V_1}{\kappa} \quad (3.5)$$

Note that since the node V_1 of transistor Q_1 is clamped by the bias current, the Early effect and back-gate effects in the phototransistor as well as Q_1 are negligible. Unlike the logarithmic receptor, the active feedback receptor should not significantly be affected by these additional effects and therefore expect the observed gain to match that predicted by equation (3.5).

The circuit [Mead90] can easily be modified to provide higher gain by adding a diode connected transistor to the photoreceptor loop as shown in Figure 3.11. The reason why this modification provides higher gain is the same as in the standard logarithmic photoreceptor. As the photocurrent increases (due to a high incident light intensity), V_2 will increase twice as fast as it did before because the gate-source voltage on Q_2 . In fact V_2 increases even faster than before because



the back-gate effect forces the gate-source voltage on Q_1 to be even higher to produce the same current as before. Due to the logarithmic response of the circuit there is no elegant way to provide an adjustable gain while preserving the circuit's function [Mead90].

Figure 3.10 Active Feedback Device

more rapidly decreases the current through Q1, than does the movement of V1. With V1 decreasing and V2 increasing, there will

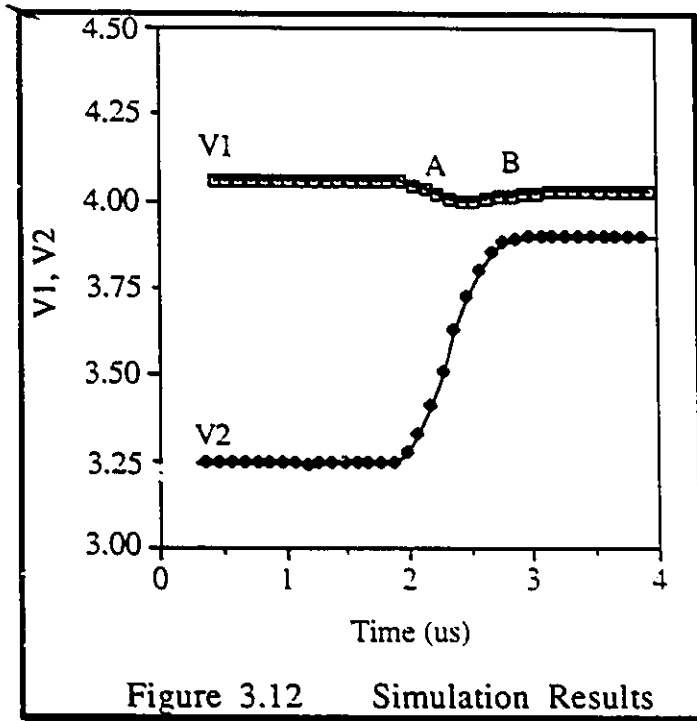


Figure 3.12 Simulation Results

be a point where the two voltages cause Q1 to accommodate the current supplied by the phototransistor. This corresponds to point A on Figure 3.12. However, V2 must continue to move towards VDD, and therefore there is a compensating effect (second order) when

V1 again rises towards its steady state value. This corresponds to point B(Figure 3.12).

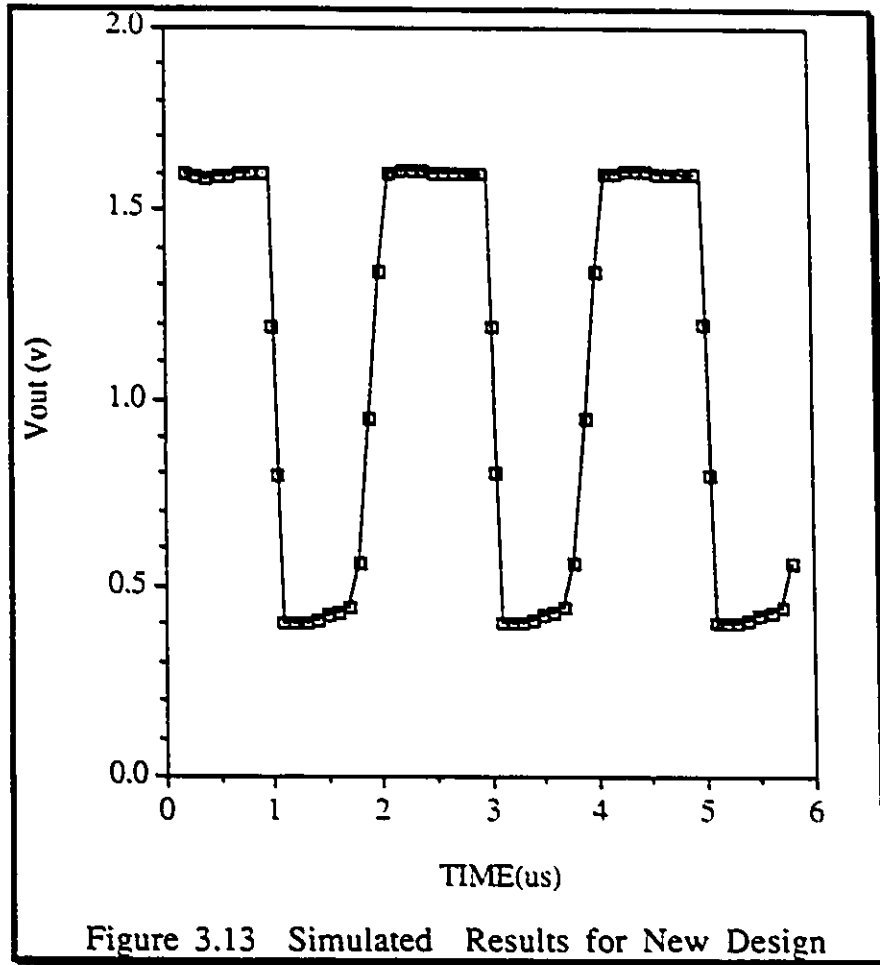
3.8.1.3 Comparison with Logarithmic Receptor

The standard logarithmic receptor has the severe handicap in that its time response is asymmetric to increases and decreases to light intensity. In fact it was shown in previous sections that the response to the removal of light is logarithmic in nature. The active feedback photoreceptor does not remove this asymmetry, since this is a property of the phototransistor. Recall that the large fall time is due

to the discharge of the large base-collector capacitance. The active feedback photoreceptor however does have the advantage that it exhibits a much faster response to the cessation of light than does the standard photoreceptor. Figure 3.13 shows the active feedback photoreceptors response to a step change of maximal intensity (maximum illumination to no illumination). The time constant of the active receptor is much smaller than that of the logarithmic receptor. This faster on-off time (especially off) improves on the logarithmic receptor design. Note the gain of the standard receptor is greater than that of the active feedback receptor. However, a simple modification to increase the gain of the active feedback receptor is given in Figure 3.11.

3.8.1.4 Conclusions

In this section, a simple modification of the standard photoreceptor is given. A photoreceptor circuit which delivers an output voltage which is logarithmically related to light intensity can be implemented using active feedback and is currently being fabricated. This design, holds the output voltage of the phototransistor, eliminating the Early effect and thus improving upon the standard logarithmic photoreceptor. The feedback employed has the property of speeding up the standard output response to changes in light intensity as predicted by the simulation results (Figure 3.13).



3.8.2 Optical Detector Using MOSFETS

Up till now the author has been discussing the fabrication of photoreceptive element using vertical BJT. However, a CMOS fabricated MOSFET in a floating p-well is optically sensitive and can be used for photoreception. It is felt that this device is feasible for implementation in Northern Telecoms 3u CMOS process.

3.8.2.1 Operation of the Device

Transistors fabricated in a p-well CMOS process have been shown to be optically sensitive when the wells are electrically floating [GaBa80]. Within the device, the photoeffect occurs primarily at the p-well and substrate p-n junction. The built in electrical field at this junction sweeps optically generated holes into the p-well. This also results in electrons being swept into the source junction in order to maintain equilibrium. The electrons can move in one of two directions: into the channel by the electric field present if the source is small (minimum dimensions) or into the substrate forming a parasitic BJT mode of operation as shown in Figure 3.14 [KiDaJoSu87].

The p-well source induced current I_{ind} in response to the accumulated holes, increases linearly with the light intensity. The current induced and the p-well potential, V_{pwell} , is given by rearranging the diode equation:

$$V_{pwell} = \frac{kT}{q} \ln \left(\frac{I_{ind}}{I_0} + 1 \right) \cong \frac{kT}{q} \ln \left(\frac{RP_{opt}}{I_0} \right) \quad (3.6)$$

where I_0 is the reverse bias saturation current (A)

R is the responsivity in (A/W)

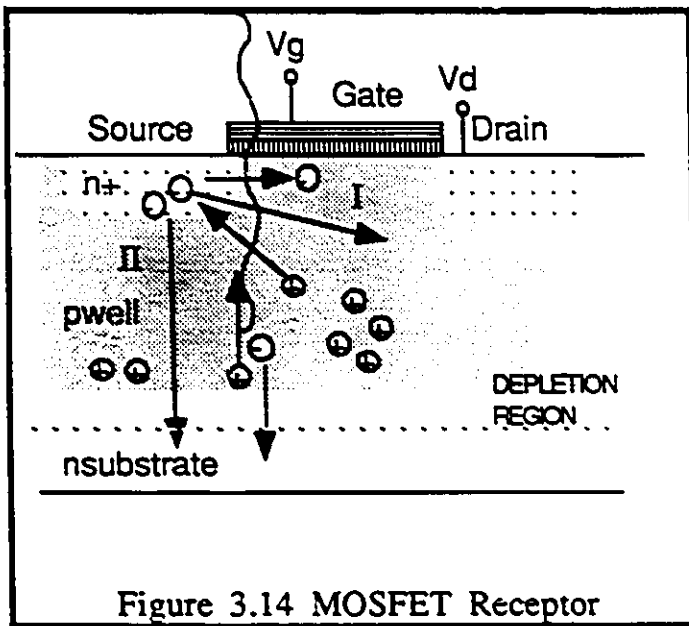
P_{opt} is the light power in W

In this expression I_{ind} and I_o are both proportional to the junction area, whereas the p-well potential is independent. As the junction area increases, more holes are generated but accumulate in a pwell volume that is larger by the same factor. Therefore, there is not net increase in the positive charge buildup of the p-well for a larger exposed area.

One of the factors that affect the threshold voltage is the body effect. An increase in the p-well potential results in a decrease in the threshold voltage according to [HoAl88]:

$$V_{th} = V_{th0} + \gamma \left(\sqrt{(2\phi_f - V_{pwell})} - \sqrt{2\phi_f} \right) \quad (3.7)$$

where γ is the bulk threshold parameter (\sqrt{V})
 ϕ represents the surface potential (V)



V_{pwell} represents the p-well potential (V)

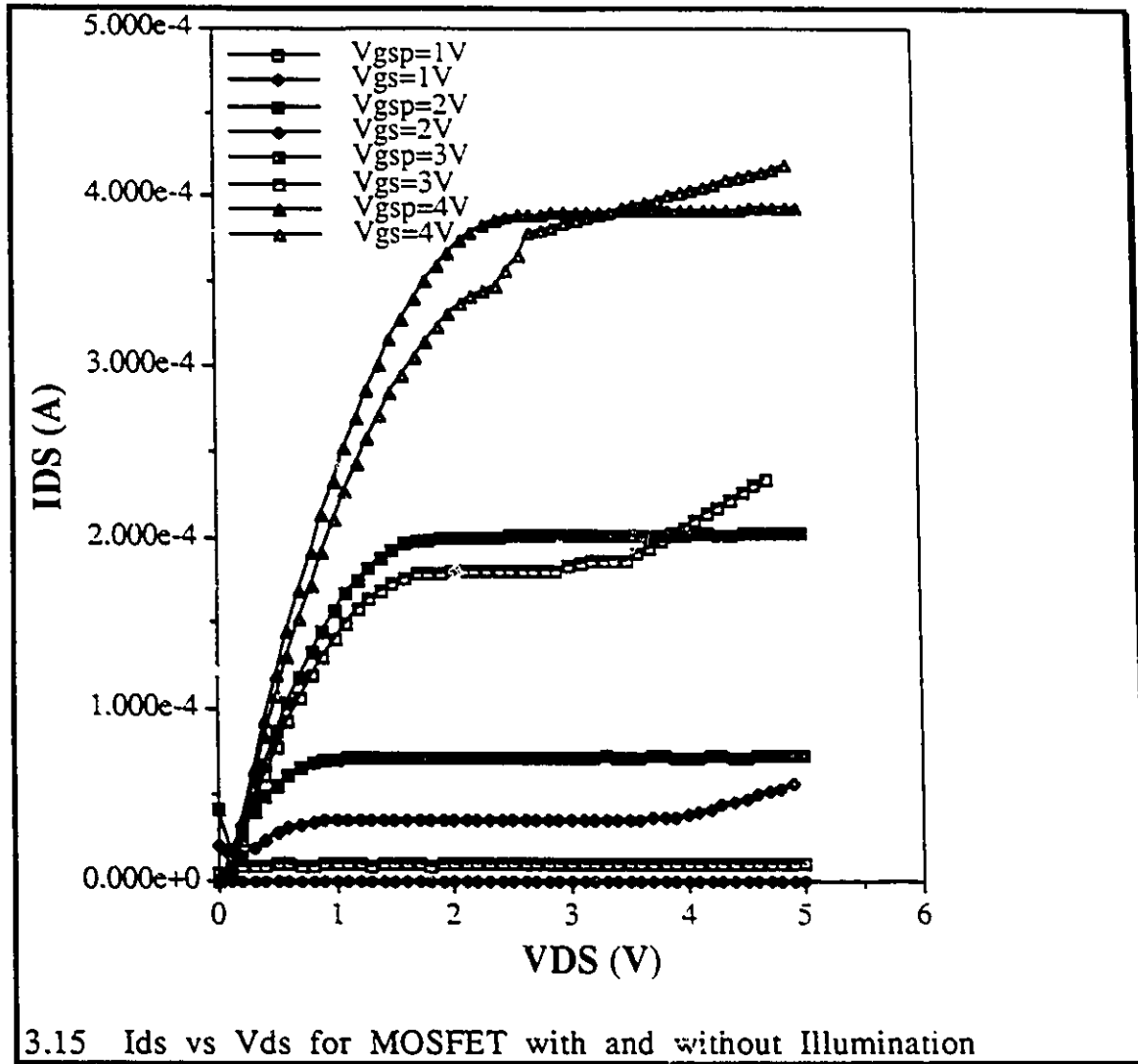
If the previous expression for V_{pwell} (equation 3.6) is substituted it can be seen the threshold voltage logarithmically decreases with increasing light intensity.

The MOSFET operating in the linear region, the drain current is related to V_{th} by the expression [HoA188]:

$$I_D = \beta [(V_{GS} - V_{TH}) V_{DS} - \frac{V_{DS}^2}{2}] \quad (3.8)$$

If the expression for the threshold voltage is substituted (equation 3.7), a relationship is formed between the drain current and the incident light power. I_D will vary approximately as $\ln(P_{opt})$ when the other variables are held constant.

For large V_{ds} ($>4V$), impact ionization occurs in the pinch off region near the drain. The extra holes formed flow across the p-well source junction and cause the p-well potential to rise, thus decreasing the threshold voltage, and increasing the channel current. A large number of excess holes also accumulate in the p-well, near the source, causing positive voltage backbiasing to occur. Figure 3.15 shows the I_{DS} versus V_{DS} for several gate voltages when the device is illuminated (V_{gsp}) and not illuminated (V_{gs}). There is a noticeable kink (Figure 3.15) in the characteristic curves as the drain current changes quickly, to indicate the presence of backbiasing.



Once the transistor is illuminated, the kinks smooth out and the p-well achieves a relatively uniform potential over the Vds range. The greatest change in drain current occurs for values of drain voltage below where the kink occurs. For values of Vds larger than the "kink" voltage, impact ionization becomes the dominant effect, and very little photo-induced change in current is present [KiDaJoSu87].

The above discussion is assuming the use of a small source area. However, if a large source area is created, a vertical npn BJT becomes dominant and this has been discussed previously.

3.8.2.2 Summary

The optical characteristics of CMOS fabricated MOSFETS have been tested and it is found that this device is appropriate for implementation in a Northern Telecom 3u CMOS process. The logarithmic relationship between threshold voltage or drain current and light intensity has been shown. These logarithmic relationships enable the device to act as a viable photodetector for practical applications.

3.9 Conclusions

This chapter has presented the reader with the background knowledge to understand the operation of the proposed photoreceptive element. A description of the electronics theory behind the element is presented such that the standard cell for the photoreceptor can be readily integrated into a VLSI design. Also presented is a SPICE simulation of the receptor circuitry. This provides the reader with values that can be verified through experimental work.

The transient response of the photoreceptor was given, in order to provide the reader with values of rise and fall times which were experimentally found. Also the dynamic range of the logarithmic relationship between output voltage and incident light intensity was experimentally determined. The presence of a peak in the spectral response was explained. The presence of the Early effect was explained, and how this effect alters the expected light intensity/current relationship.

Also presented is the design of a photoreceptor with active feedback. This implementation reduces the effect of the Early voltage, as well as improving the response to the cessation of light. The optical properties of MOSFETs were presented and it was shown that this device has a logarithmic relationship between the drain current and light intensity.

In all three structures (the logarithmic photoreceptor, the active feedback photoreceptor and the floating well MOSFET), the feasibility of implementation in Northern Telecoms 3μ CMOS process is demonstrated.

IV. An Optically Coupled Differential Current Mode Neuron for Neural Networks

4.1 Introduction

Over the last decade a great deal of research effort has gone into not only into the operation of the human brain, but also in the VLSI implementation of the brain's architecture into silicon. An artificial network is comprised of three distinct parts as shown in Figure 1. These parts consist of the weighted input connection, the summation function, and a threshold function, which depends on the states of the input elements and the connection strengths, or weights. Although each piece corresponds to a portion of the biological neuron, it can clearly be shown that the biological neuron is far more complex than the artificial neuron. The operation of this neuron can be mathematically expressed as

$$I_{out} = f \left(\sum_{i=0}^{n-1} S_i W_i \right) \quad (4.1)$$

where I_{out} is the output of the network, f is the thresholding function, the $\{S_i\}$ are the input strengths, $\{W_i\}$ are the input weights, and n is the total number of inputs.

Although a great deal of research is presently being invested into the VLSI implementation of neural networks, the problems of the large number of input and output pins required for the network are constantly arising [Isma89]. As the complexity of a neural network is increased by adding additional nodes, the number of weights required

(for a fully interconnected network) required increases by a factor of 2. This problem has been handled differently by different research groups producing a wide variety of solutions. One of the better implementations includes EEPROM technology for the storage of the weights on chip[HiGo89].

The author has been charged with the implementation of an optically coupled neural network. The inputs provide the basis for sensor fusion and an optically programmable weighting capability (for off chip learning).

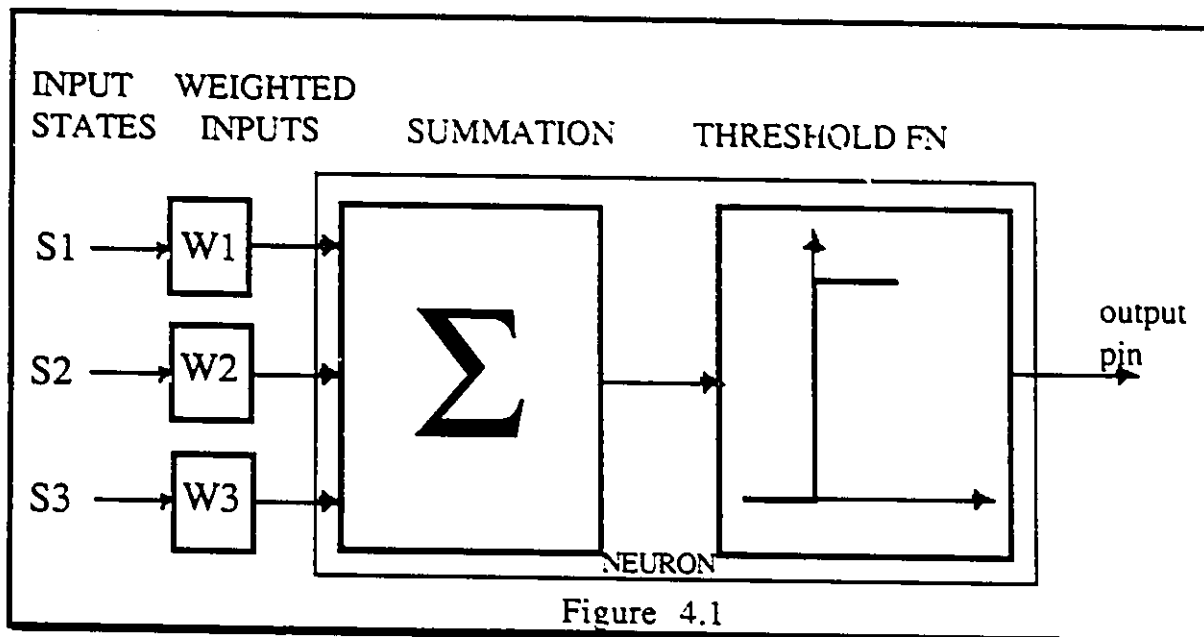
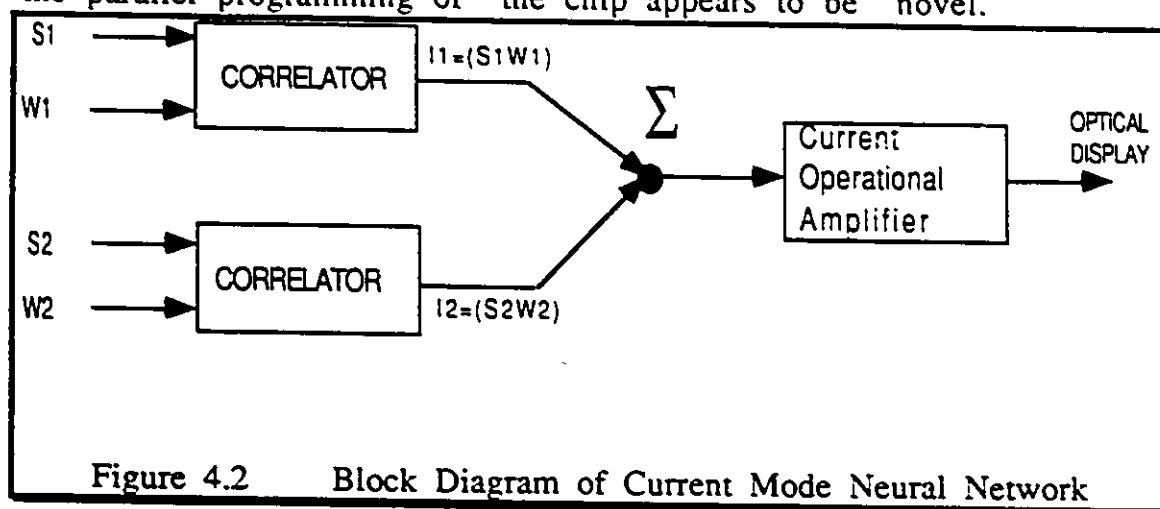


Figure 4.1

A previous research project at our labs included the realization of a photoreceptor integrated with a floating resistor, which provides weights ranging from 90K to 180K [BeJuMi89]. This approach takes the output of a photoreceptor (implemented in CMOS as a series parasitic BJT and a sub-threshold MOS chain) as a voltage which is used to control a floating resistor circuit. The resistor then performed the functions of the synapse connected to a neuron. The implementation discussed in this paper uses current as opposed to voltage to drive the

network. Current is far better suited to an analog neural network implementation due to the inherent property that current is summed at a node, thus eliminating the use of complex voltage summation circuitry. Recall that summation is one of the major subblocks in Figure 4.1. Current signals do not experience the "drops" that voltage signals do, as the signal propagates through the network. Work in programmable current mode neural networks for analog VLSI is being done by Ismail [Isma89]. Using the current mode architecture, a new block diagram of a neural network is shown in Figure 4.2.

This paper demonstrates the use of an optoelectronic device to provide inputs and weights to a current mode analog neuron. The idea of inputting data to an integrated circuit by light is not new, CCD cameras rely on this principle of operation, but the use of light beams to perform the parallel programming of the chip appears to be novel.



The author uses a "correlator" circuit [Mead90] in place of a true analog multiplier to find the product of the optical inputs and optical weights. This product (represented as current) is then summed and is fed into a current operational amplifier which performs the thresholding function on the current. The final output nodes of the

network will be connected to an optical display (as discussed in Appendix A) which will eliminate the need for output pins (internal nodes will be connected using metal). It should be noted that this display can be implemented regardless of the driving mechanism. It can be seen from Figure 4.2 that the operations performed in the block diagram of a voltage driven network can be more easily performed in a current driven network by making use of the properties of current that were discussed.

The novel implementation of the multiplier as a correlator circuit reduces by as much as 50% the number of devices required for an analog multiplier [HoAl89]. For analog circuitry, this reduction translates into a larger number of synapses in the same area.

The current mode architecture contains transistors operating in the subthreshold region. For sufficiently low gate voltages the drain current of a transistor is given by [Mead89]:

$$I = I_0 e^{\kappa V_g - V_s} (1 - e^{-V_{ds}}) \quad (4.2)$$

where I is the transistor current
 κ is the field correlation factor

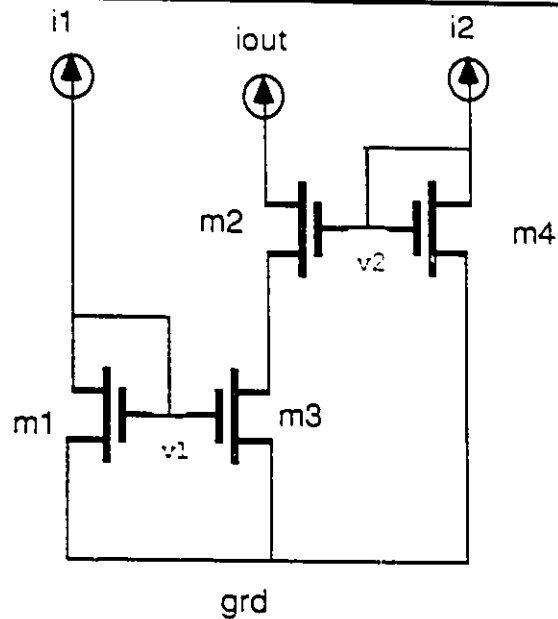
When designing in subthreshold it is important to realize that this region of operation both limits the power consumption of the chip as well as provides exponential characteristics for the transistors. There does however exist the disadvantage of poor matching characteristics between transistors in subthreshold.

The circuits used to implement the blocks of the current mode neural network block diagram will now be presented. In each

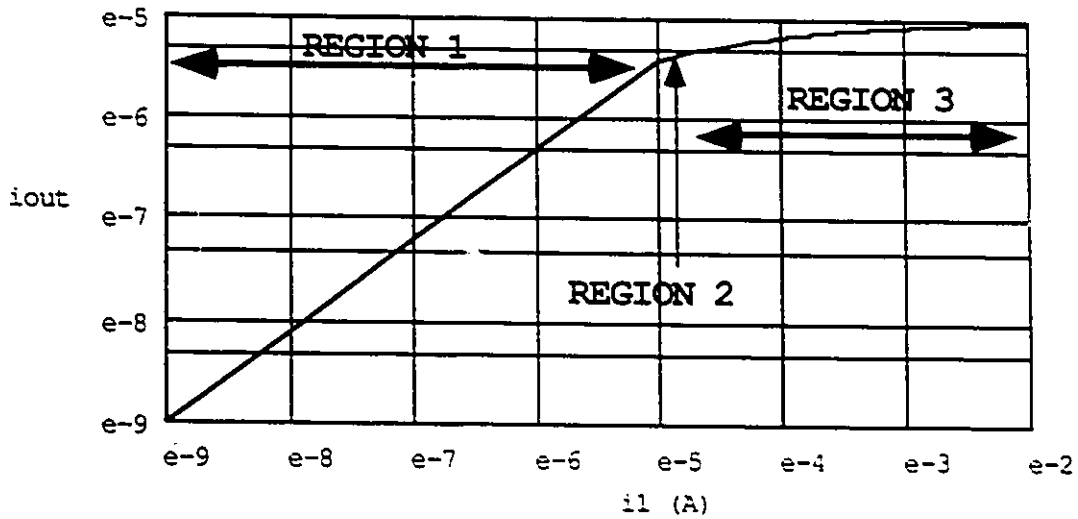
discussion, starting with the correlator, the simulation results as well as a discussion of the application of the circuit is given.

4.2 Approximate Multiplication by Correlation

This circuit replaces the usual synapse analog multiplier and is based on the concept of MOSFETS used as binary switches. The correlator circuit is shown in Figure 4.3a and consists of 4 transistors operating in the subthreshold region. The operation of the correlator is based on the operation of the transistors acting as binary switches; used in the on position to bring in the weights. A condition that is required is that the input to the neuron should be much larger than the analog weighting current from the photoreceptors. This is the case since the current from our photoreceptors is less than 0.1uA, whereas the current from the output of the previous layer is an order of magnitude higher.



(a)



(b)

Figure 4.3 (a) Correlator Circuit (b) Spice Simulations

By simple analysis the transfer function of the circuit can be found as follows:

Assuming that M2 is in saturation (since $I_2 \gg I_1$) produces from equation (2):

$$I_{m2} = I_0 e^{kV_2} e^{-V_n} \quad (4.3)$$

Assuming M3 is not saturated producing:

$$I_{m3} = I_{03} e^{(kV_2)} (1 - e^{-V_n}) \quad (4.4)$$

Expressions for I_2 and I_1 can be shown as:

$$I_2 = I_0 e^{kV_2} \quad (4.5)$$

$$I_1 = I_0 e^{kV_1} \quad (4.6)$$

Since transistors m2 and m3 are in series, equating currents from (4.3) and (4.4) produces:

$$I_{out} = I_{02} e^{kV_2} e^{-V_n} = I_{03} e^{(kV_2)} (1 - e^{-V_n}) \quad (4.7)$$

Finding an expression from equating (4.5) and the first part of (4.7) and solving produces

$$e^{-V_n} = \frac{I_{out}}{I_2} \quad (4.8)$$

Substituting this expression into latter part of (4.7) and comparing with (4.6) produces

$$I_{out} = I_1 \left(1 - \frac{I_{out}}{I_2} \right) \quad (4.9)$$

Producing [Mead90]:

$$I_{out} = \frac{I_1 I_2}{I_1 + I_2} \quad (4.10)$$

The circuit can be considered to operate in three regions as shown in Table 1.

Region	Inputs	Outputs
1	$I_1 \ll I_2$	$I_{out} \cong I_1$
2	$I_1 \cong I_2$	$I_{out} \cong 0.5 I_1$
3	$I_1 \gg I_2$	$I_{out} \cong I_2$

Table 1

The circuit is forced to operate only in regions 1 and 3. In region 3 I_1 is 0A producing an output of 0A whereas in region 1, $I_{out} = I_2$. Basically it is this latter region that represents the linear part of the simulation curve (Figure 4.2b), where the analog weights are being switched in. The correlator discussed above performs the "multiplication" for two positive currents: the fourquadrant correlator will allow the "multiplication" of a positive or negative input with a positive or negative weight (from either the positive or negative photoreceptor) as shown in Figure 4.4.

Figure 4.5 shows the four quadrant correlator which consists of four , one quadrant correlators. Clearly the two branches on the left hand side provide correlation for the positive output current- produced by photocurrents and neuron current of the same sign. The two right hand side branches provide sink current, produced by photocurrents

and neuron currents of different signs [Chan90]. The positive current (from the left hand side) is then mirrored into the output node.

Although this four quadrant correlator provides current outputs for any of four possible cases usually only positive or negative weights are required. Therefore branches can be removed quite easily without changing the performance of the circuit. The ease in customization makes this subcircuit very powerful.

The output current of the four quadrant correlator can be expressed as a function of the different input currents by writing KCL at node Out:

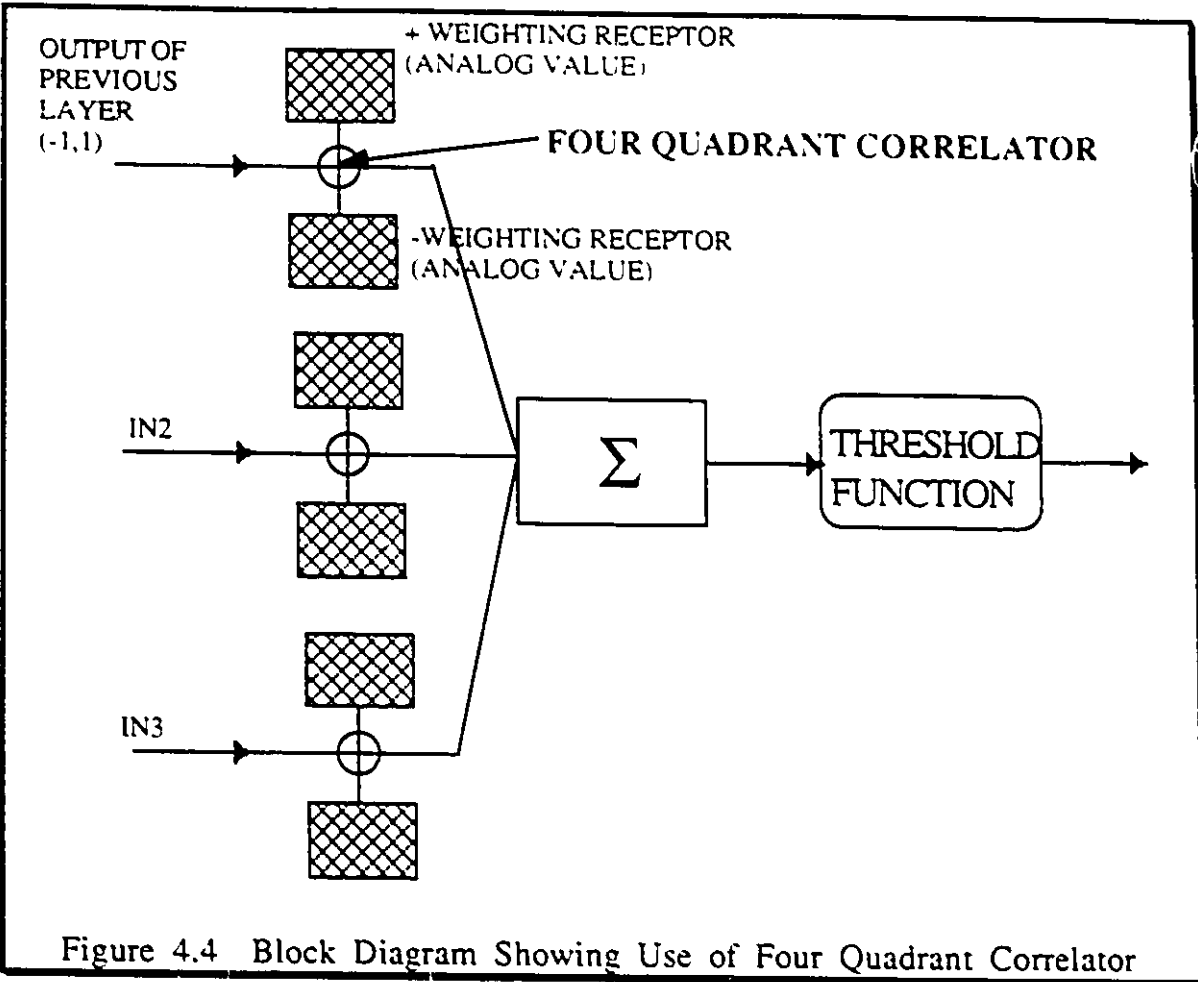
$$I_{out} = I_{np} \left[\frac{1}{1 + \frac{I_{np}}{I_{pp}}} - \frac{1}{1 + \frac{I_{np}}{I_{pn}}} \right] + I_{nn} \left[\frac{1}{1 + \frac{I_{nn}}{I_{pn}}} - \frac{1}{1 + \frac{I_{nn}}{I_{pp}}} \right] \quad (11)$$

where I_{nn} , I_{np} are negative and positive input current respectively.

I_{pn} , I_{pp} are negative and positive weighting photo currents respectively.

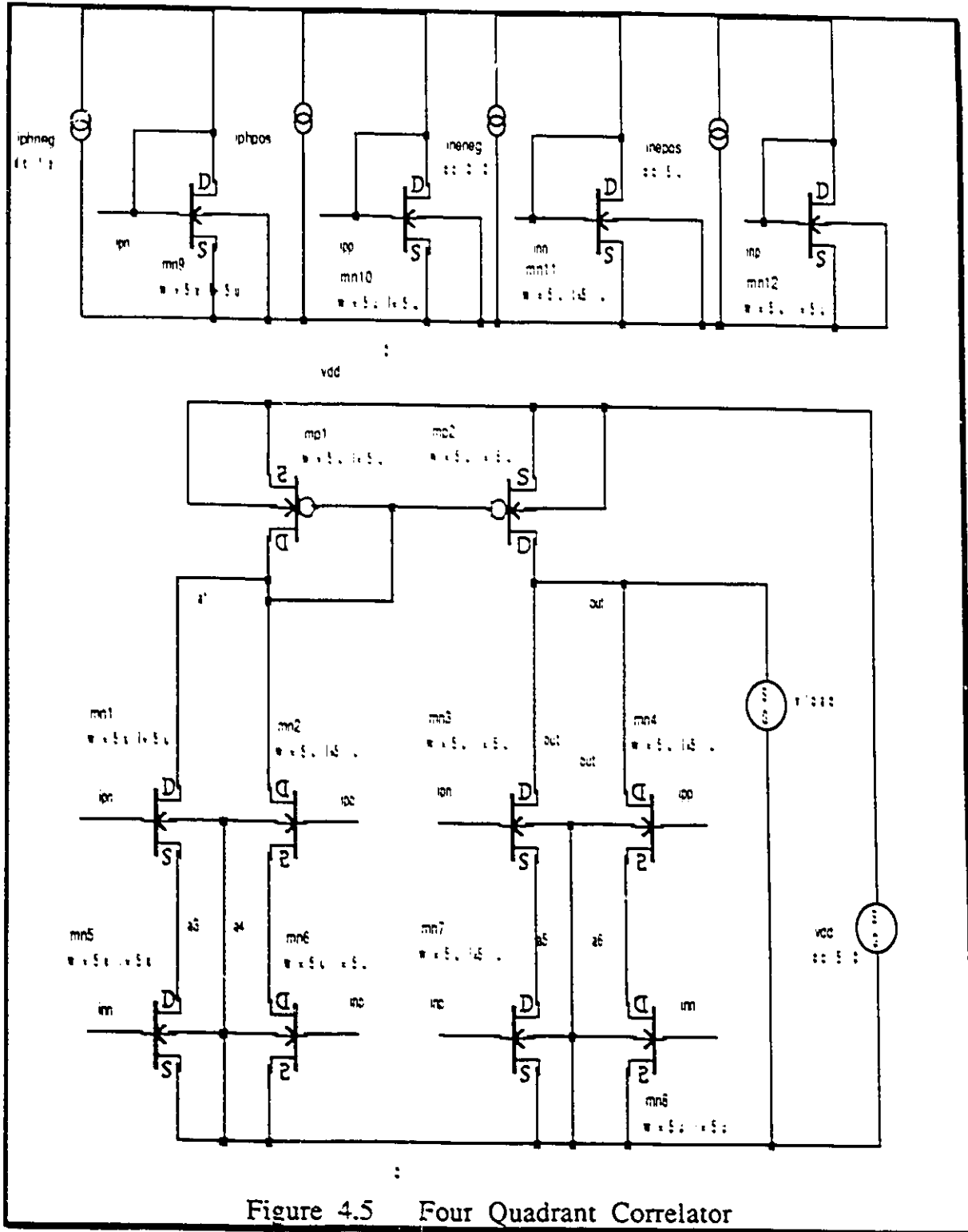
It can be seen that only one current represents both the excitatory or inhibitory synaptic weights. The distinctive feature is achieved by the current mode architecture which has been implemented.

The product of the "digital" input current and the analog weighting current can now be calculated. As shown in the block diagrams above, this must now be thresholded, as will be discussed in the following section.



4.3 Thresholding the Weighted Input

The current output of the correlator circuit is then connected to a current mode operational amplifier which provides the linear, thresholded, neuron characteristic. This characteristic is implemented using the fact that currents into the circuit are mirrored using the



appropriate type of transistors (n channel MOSFETS mirror sinking currents, p channel MOSFETS mirror sourcing current) into a common node. Using the fact that currents at a node are summed (with signs) at a node, I_{out} represents the thresholded signal.

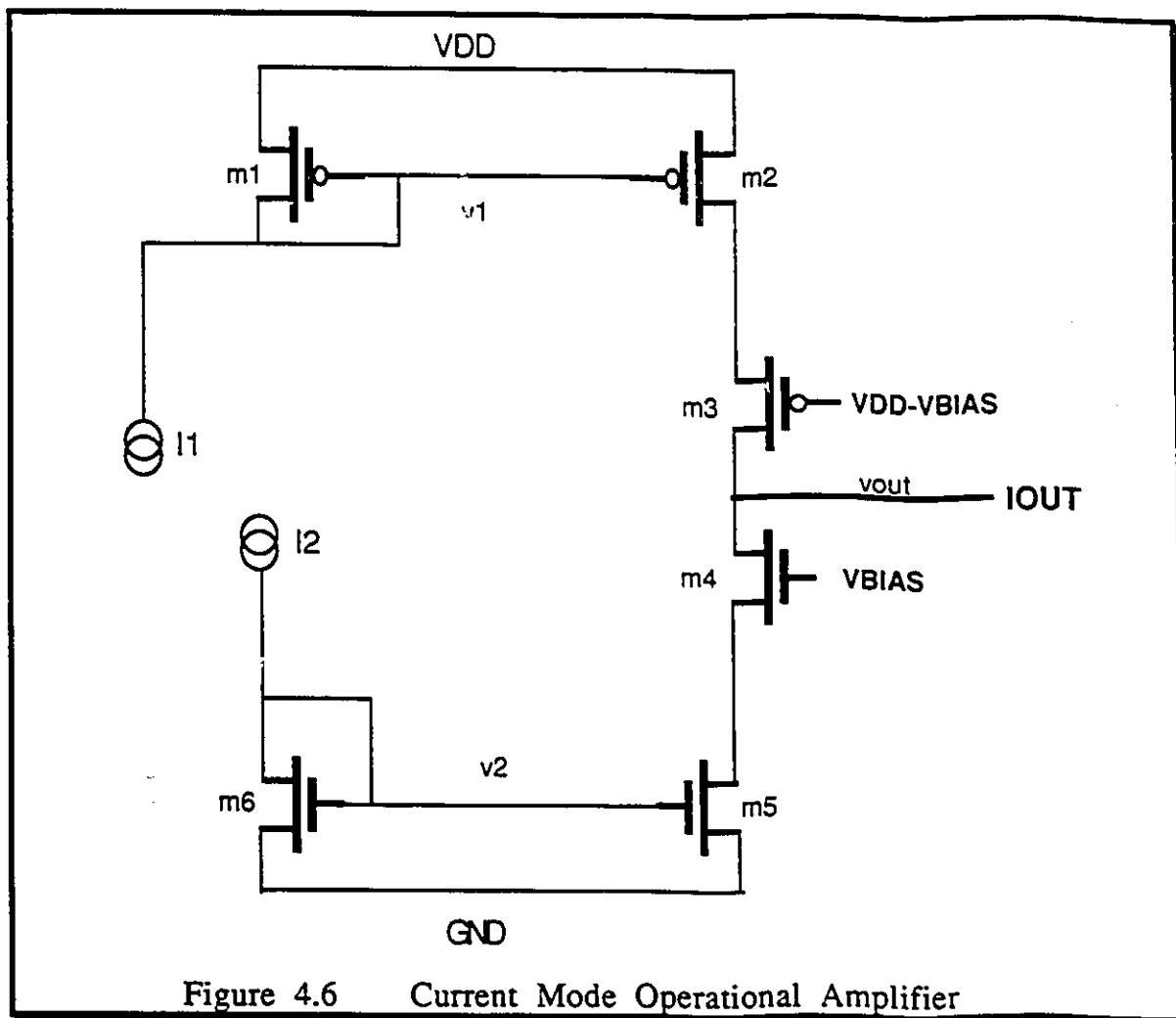


Figure 4.6 Current Mode Operational Amplifier

The output current of the current operational amplifier can be expressed (writing KCL at node I_{out}) as :

$$I_{out} = I_{m3} - I_{m4} \quad (4.12a)$$

which can be expanded to

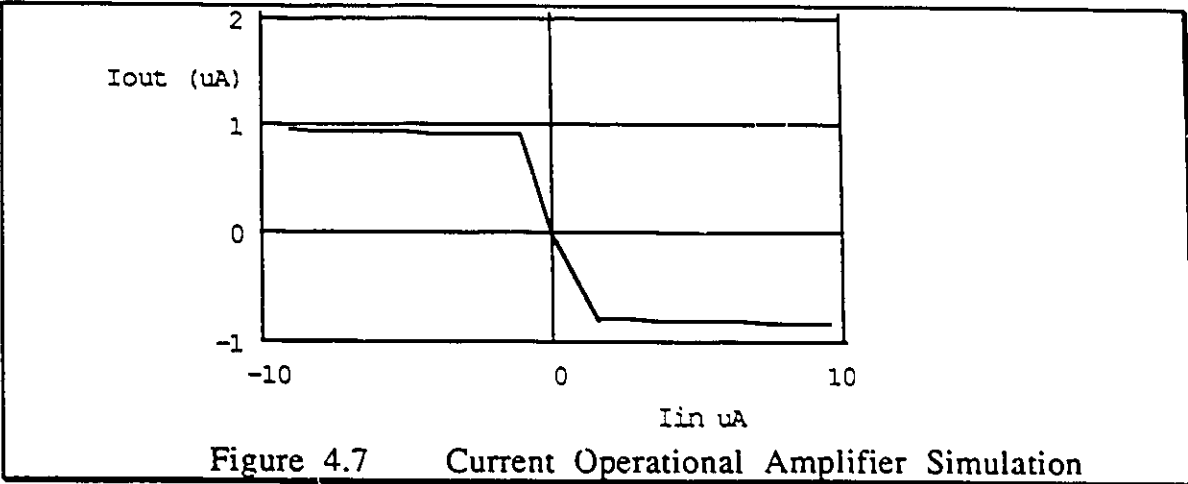
$$I_{out} = I_{m2} e^{\kappa(V_{dd}-V_{bias})-V_1} (1 - e^{-(V_{out}-V_1)}) - I_{m4} e^{\kappa(V_{bias})-V_1} (1 - e^{-(V_{out}-V_2)})$$

(12b)

where I_{m2} and I_{m4} represent the mirrored input currents I_1 and I_2 respectively as shown in Figure 7. The figure shows that the amplifier saturates at $-1\mu A$ and $1\mu A$ and has a linear region between $-2\mu A < I_{in} < 2\mu A$. The saturation level of this current amplifier is determined by the gate voltages applied to m_3 and m_4 , V_{bias} and $V_{dd}-V_{bias}$ respectively. The slope of the linear region (current gain) of the amplifier is determined by $[(w/l)_5 / (w/l)_6]$. The offset shown in Figure 7 is caused by transistor mismatches due to the operation in subthreshold region.

Clearly for different applications the gain will have to be determined to achieve the correct gain. The amplifier provides a variety of thresholding function; the linear function shown will tend toward a hard limiting function when the amplifier gain is sufficiently high. The ability to provide a linear thresholding function makes this circuit suitable for neurons to be used in neural associative memories [Mead89]. The thresholding function with sufficiently high gain can be implemented in networks which are to be used in character recognition.

The implemented structure clearly uses the properties of current mode circuitry to full advantage in using only 6 transistors. This smaller circuitry clearly allows the implementation of more neurons in the same area: ideal for large system applications where thousands of neurons are not uncommon.



The final part of the network consists of directional sensing circuitry. This circuitry is used to separate the positive and negative currents in order to maintain the differential architecture. The entire network is shown in Figure 4.8 and the spice simulation for positive correlation is shown in Figure 4.9. The results clearly show the output current (at the appropriate differential node) is increasing for the case where I_{pp} and I_{np} are correlated. The current passing through the other node does not change.

Now that the thresholded signal is found, the output will go to a postprocessed liquid crystal display (Appendix A). It should be noted that the output of this differential current mode network can just as easily be sent to an output pin on the I.C. I feel an investigation into new optical I/O methods for neural networks is well merited.

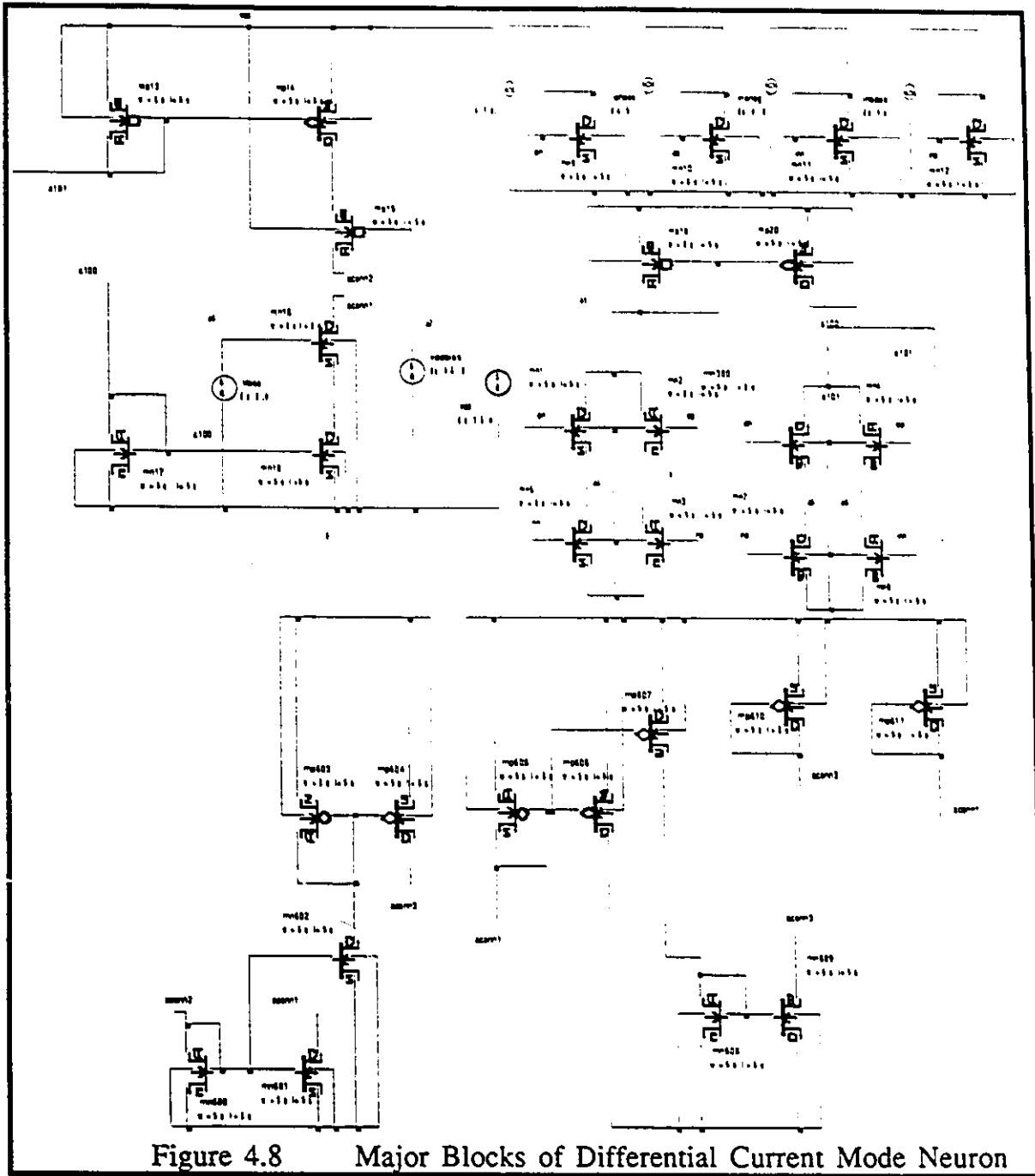
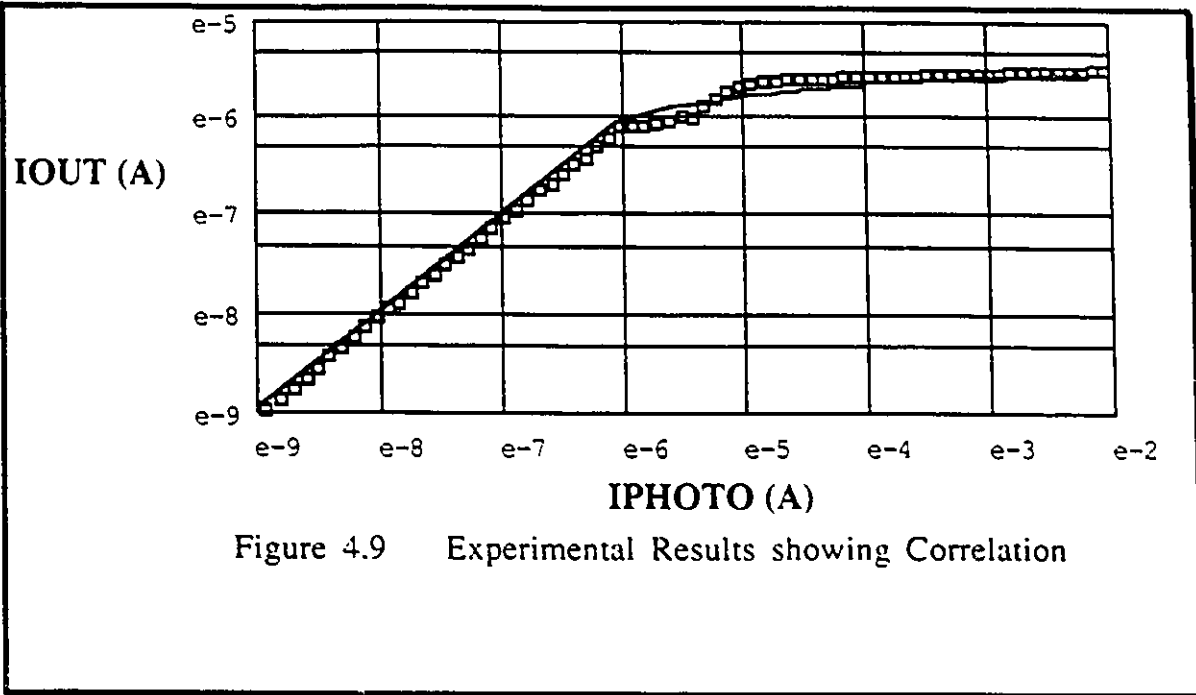


Figure 4.8 Major Blocks of Differential Current Mode Neuron



V. CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions and Recommendations

The research described in this thesis deals with CMOS photosensitive devices suitable for use in an optically coupled neural network. The inputs of this network provide the basis for the sensor fusion and an optically programmable weighting capability (for off chip learning). The constraint imposed was that this neural network must be implemented in Northern Telecoms 3u process. A set of conclusions and recommendations from the work done is best presented in point form:

1) A thorough literature search resulted in a first order model of the phototransistor. This consisted of a reverse biased photodiode in parallel with the base collector of the BJT (Section 2.1.3). This model is valuable in its ability to explain some of the experimental data obtained later in the course of the work. This model consisted of a reverse biased photodiode in parallel with the base collector of the BJT. Using this model a photoreceptive element was implemented in CMOS. The implemented device consists of a vertical npn BJT. The collector is formed by the n type substrate. The base is formed by the 100μ by 100μ layer p-well and the emitter is formed by a 15μ by 15μ layer of n plus. This structure is shown in Appendix I (Figure 2.10 shows a cross section).

2) The second stage of the development of the optically coupled neural network consisted of characterizing the operation of the photoreceptive device. The device is responsive up to 4.1 mW/cm^2 producing a saturation voltage of 4.8V (Figure 3.1). The optical dynamic range of the photoreceptive device was experimentally

determined using a set of neutral density filters. The output voltage of the device is linear over five orders of magnitude of incident light intensity. The response of the photoreceptor to cessation of light was experimentally determined using a square wave input to a He-Ne laser (as shown in Appendix IV). The 100 μ by 100 μ photoreceptor exhibited an exponential decrease in voltage upon the cessation of incident light. The photoreceptor took 10 μ s to return to a dark output voltage. When designing the photoreceptor, keep the size of the photoreceptor no larger than the 100 μ X100 μ size. This is due to the fact that the larger the photoreceptor area, the longer the response time. The response time can be calculated for various size photoreceptors as shown Section 2.1.2.5. The device exhibited a great deal of crosstalk between adjacent photoreceptors. If a design requires the implementation of an array or grid of photoreceptors, it is crucial to place minimum dimension p-well between adjacent devices. The crosstalk without this design consideration will certainly corrupt your data. Large size photoreceptors tend to have large voltage variations in their p-well. The implementation of a large number of contacts (make them minimum size) must be as shown in Appendix I.

3) As the author was charged with the implementation of the most robust photoreceptor, it became important to enhance the performance of the device. The effects of the Early voltage will cause a variation in the light intensity/collector current relationship (when the output voltage is changed). When the collector emitter voltage of the BJT is being driven over 3V, use the circuit presented in Figure 3.10. Not only will this device enhance the performance of the device, it will certainly speed it up if implemented correctly. The speed of the circuit will improve the response time of the standard

photoreceptor to 1 μ s (a ten fold improvement over the conventional photoreceptor). When laying the circuit, keep all devices minimum dimensions. It is crucial that a p-well is placed around the BJT to prevent latch-up problems associated with Q1 of Figure 3.10. The BJT should include the minimum dimension contacts in order to maintain a constant voltage in the p-well.

4) The author has been successful at implementing the photoreceptive devices in an optically coupled neural architecture especially designed to make use of the possibility of optical programming. The photoreceptive devices have been successfully used as inputs providing the basis for sensor fusion. The ability to optically program the network (off-chip learning) has been demonstrated. The author has made use of a switched gate circuit which multiplies the input to the input current by the optical programming current. This reduces the number of transistors required for the implementation of the synaptic function by 50%. The author found it was best to keep the transistors minimum dimensions, thus reducing the area of the circuit while maintaining operation. This reduction becomes more valuable in the implementation of a four quadrant switched gate circuit. This circuit provides for positive as well as negative weighting currents. The neural network was tested and was in good agreement with simulation results as shown in Figure 4.9.

References

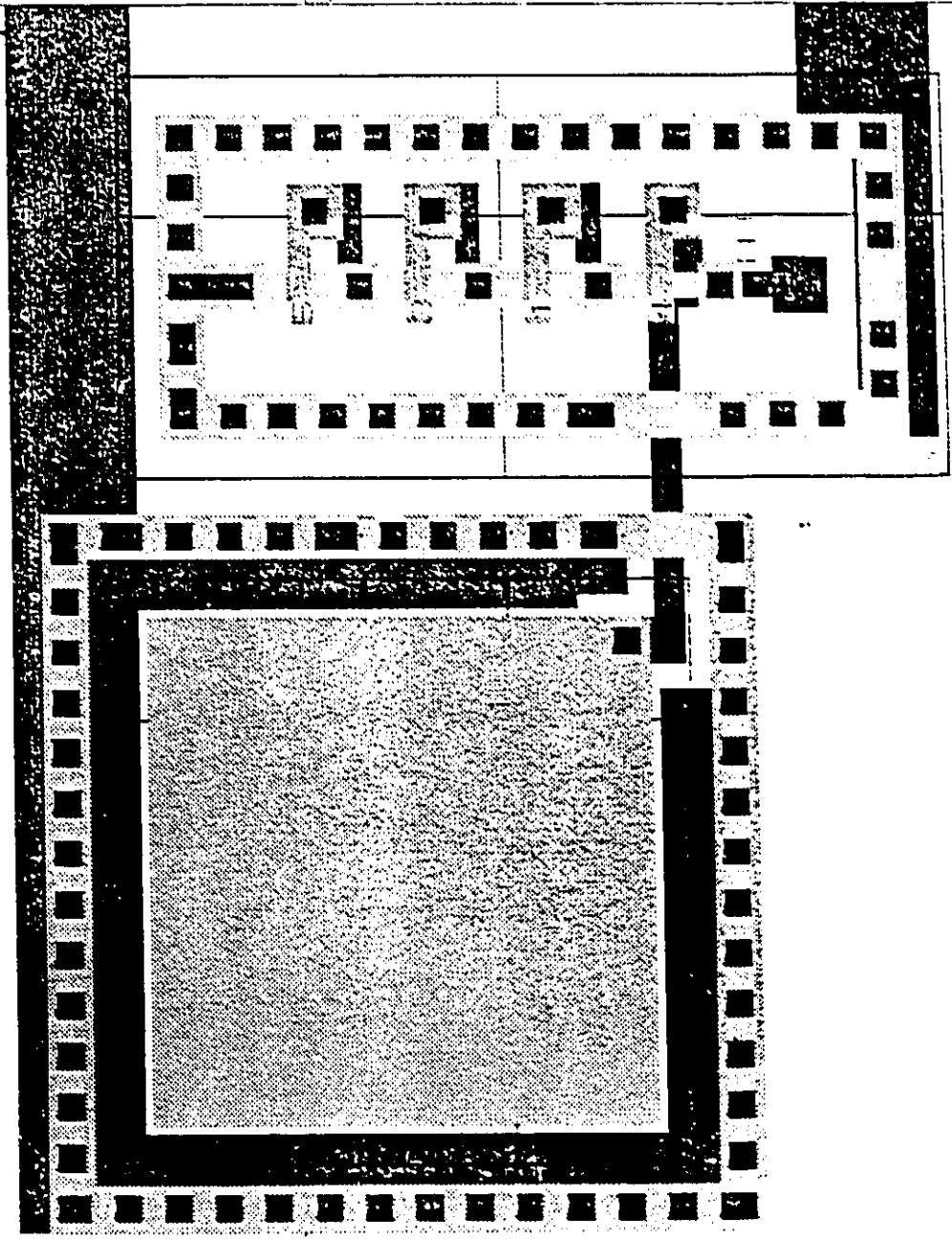
- [AnMa88] P. Antognetti and G. Massobrio, Semiconductor Device Modelling with SPICE, New York, McGraw Hill, 1988
- [CaTo68] D. Callahan and J. Torley, "Performance Characteristics of Pulsed Phototransistor Structures Under Various Conditions", IEEE Transactions on Electron Devices, ED-15, April 1968, pg248-258
- [Chan89] A. Chandna, Computational Neural Science Lab#11
- [ChJuMi89] A. Chandna, G. Jullien, W.C. Miller, "Opto-Programmable Neural Networks: An Initial Study", Proceedings of the Canadian Conference on VLSI, Ottawa Canada, 1989
- [DeWi87] J.S. Denker, B.S. Winter, "Network Generality, Training, Required and Precision Required,"IEEE Conference on Neural Information Processing, American Institute of Physics (AIP Press, New York), 1987
- [GaLi68] P. Gary and J. Linville, "A Planar Silicon Photosensor with an Optimal Spectral Response for Detecting Printed Material", IEEE Transactions on Electron Devices, ED-15, January 1968, pg 30-39
- [GrVe89] H.P. Graf, P. de Vegvar, "A CMOS Associative Memory Chip Based on Neural Networks", ISSCC Digest Technical Papers, 1989, pp304-305
- [HaNe76] W. Hayt, G. Neudeck, Electronic Circuit Analysis and Design, Houghton Mifflin, 1976
- [HiGo89] V. Hilleringmann and K. Goser, "CMOS Circuits Integrated on Laser Recrystallized Silicon Films", Proceedings URSI, University of Erlanger, September 1989, pp423-426
- [HoAl88] P. Allen, and D. Holdberg, CMOS Analog Circuit Design, HRW Press, Toronto, 1988
- [Isma89] M. Ismail, VLSI Implementation of Neural Systems, Kluwer Academic Press, New York, 1989, ch3.2, pg 27-4,207-227

- [JoLi89] R. Joy and J. Linville, "Phototransistor Operation in the Charge Storage Mode", IEEE Transactions on Electron Devices, ED-15, April 1968,pg 237-248
- [KiDaJoSu87] S. Kirkish, J. Daly, L. Jou, S. Su, "Optical Characteristics of CMOS Fabricated MOSFETS", IEEE Journal of Solid State Circuits, April 1987, pg 299-301
- [Kova87] K. Kovaly, Neurocomputing: The Technology, the Players, the Potential, Technical Insight Inc., Fort Erie N.J., 1987
- [McVaSi89] D. McKnight, D. Vass, R. Sillitto, "Development of a Spatial Light Modulator: A randomly adressed liquid crystal over NMOS array",Applied Optics, pg 4757-4761
- [Mead87] C. Mead, "A Sensitive Electronic Photoreceptor", Proceedings of the 1987 Chapel Hill Conference on VLSI, pg 450-454
- [Mead89] C. Mead, Analog VLSI and Neural Systems, Addison-Wesley, New York, 1989
- [Mead90] C. Mead et al, Characterizing the Photoreceptor CNS Lab17-Caltech, 1990
- [MoChZi71] F. Moneda, E. Chenette, and A. Ziel, "Noise in Phototransistors", IEEE Transactions on Electrons Devices, ED-15, June 1971
- [MuSm88] A. Murray, A. Smith, "A Synchronous VLSI Neural Network Using Pulse Stream Arithmetic", IEEE Journal of Solid State Circuits, June 1988, pg688-697
- [RuMc86] D.E. Rumelhard and J.L. McClland, Parallel Distributed Processing : Explorations in the Microstructure of Cognition, MIT Press, Cambridge, Mass., V1-2,1986
- [ScHoHu89] D. Schwartz, R. Howard, W. Hubbard, "A Programmable Analog Neural Network", IEEE Journal of Solid State Circuits, pg 313-318, April 1989
- [SeSm87] A. Sedra, K. Smith, MicroElectronic Circuits, Holt, Reinhart and Winston, Toronto, 1987

- [Shep89] G. Shepard, Synaptic Organization of the Brain, John Wiley and Sons, New York, 1989
- [SiEmMe85] M Sivilotti, M.R. Emmering and C.A. Mead, "A Novel Associative Memory using Collective Computations", Proceedings of the Chapel Hill Conference on VLSI, 1985, pg 329-342
- [[SiEmMe86] M Sivilotti, M.R. Emmering and C.A. Mead, "VLSI Architectures for Implementation of Neural Networks", Proceedings of the AIP Conference of Neural Networks for Computing, Snowbird, 1987, pg 408-413
- [Stree80] B. Streetman, Solid State Electronic Devices, Kluwer Academic Press, Hingham, M.A., 1980
- [Sze81] S. Sze, Physics of Semiconductor Devices, Wiley Interscience, New York,
- [TrPaVe89] P. Treleaven, M. Pacheco, M. Vellesco, "VLSI Architectures for Neural Networks", IEEE Micro, December 1989, pg 8-25

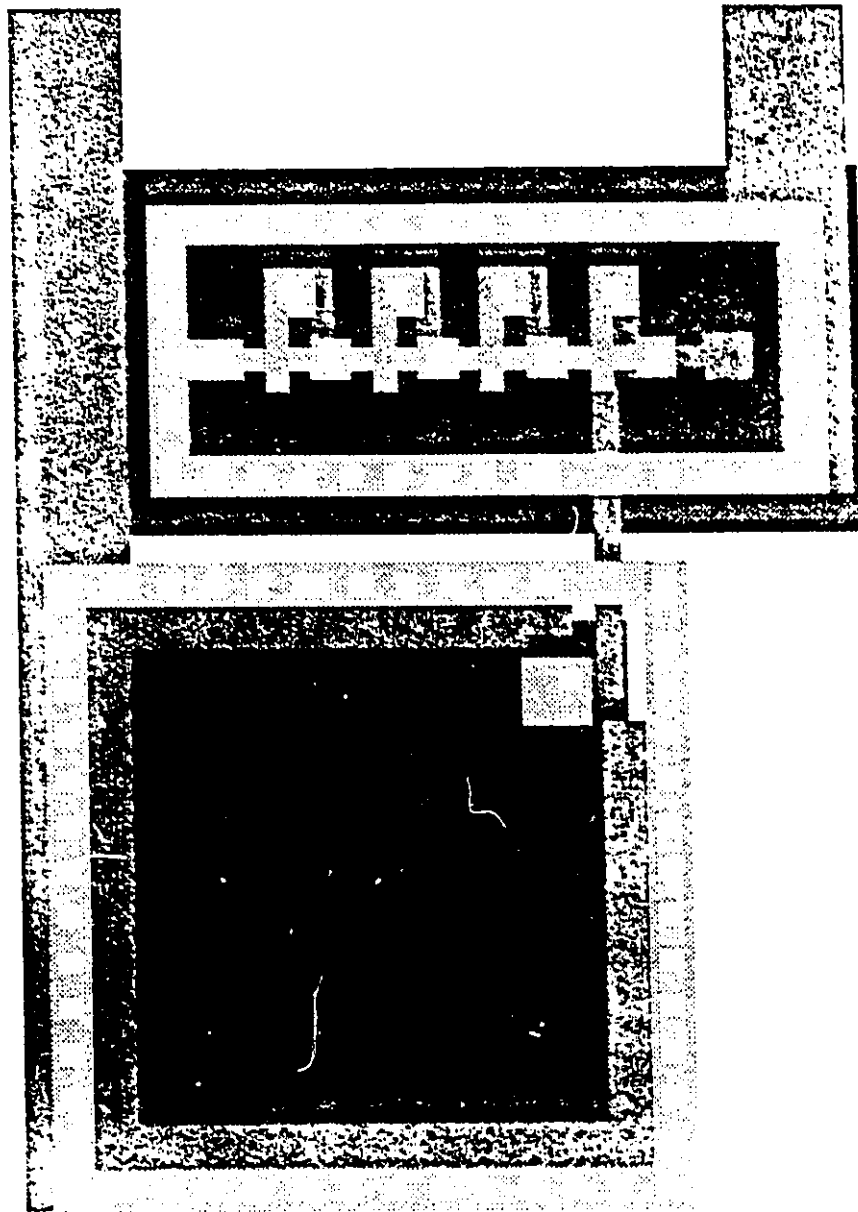
APPENDIX I
PLOT OF THE PHOTORECEPTIVE DEVICE

1. ...
 2. ...
 3. ...
 4. ...
 5. ...
 6. ...
 7. ...
 8. ...
 9. ...
 10. ...
 11. ...
 12. ...
 13. ...
 14. ...
 15. ...
 16. ...
 17. ...
 18. ...
 19. ...
 20. ...
 21. ...
 22. ...
 23. ...
 24. ...
 25. ...
 26. ...
 27. ...
 28. ...
 29. ...
 30. ...
 31. ...
 32. ...
 33. ...
 34. ...
 35. ...
 36. ...
 37. ...
 38. ...
 39. ...
 40. ...
 41. ...
 42. ...
 43. ...
 44. ...
 45. ...
 46. ...
 47. ...
 48. ...
 49. ...
 50. ...
 51. ...
 52. ...
 53. ...
 54. ...
 55. ...
 56. ...
 57. ...
 58. ...
 59. ...
 60. ...
 61. ...
 62. ...
 63. ...
 64. ...
 65. ...
 66. ...
 67. ...
 68. ...
 69. ...
 70. ...
 71. ...
 72. ...
 73. ...
 74. ...
 75. ...
 76. ...
 77. ...
 78. ...
 79. ...
 80. ...
 81. ...
 82. ...
 83. ...
 84. ...
 85. ...
 86. ...
 87. ...
 88. ...
 89. ...
 90. ...
 91. ...
 92. ...
 93. ...
 94. ...
 95. ...
 96. ...
 97. ...
 98. ...
 99. ...
 100. ...



100-404010-100

101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200



APPENDIX II
THEORETICAL BASIS OF A LIQUID CRYSTAL DISPLAY

POSTPROCESSING OF A LIQUID CRYSTAL DISPLAY

Liquid crystal research dates back to 1888 , but it was not until only recently that practical applications were realized. With the discovery of the display applications within the last five years, there has been a tremendous increase in the research and development directed at new display products.

Liquid crystal displays (LCD), are light modifiers, or light scatterers as opposed to being self-luminescent. Their attraction stems from characteristics such as viewability in high light ambients, their low cost, and their low power drain.

The most practical display phenomenon, is the electro-optic effect called dynamic scattering. Since the liquid crystal field is rapidly changing, dynamic scattering may not retain its position five years from now.

The reason for the fabrication of a liquid crystal light valve is to provide an optical device which can be postprocessed on an integrated circuit. The converged values of the network will be used to drive the liquid crystal device, thus eliminating the need for output pins. The following sections provide an overview to the chemical and physical properties exhibited by liquid crystals. A discussion of the electro-optical mechanisms that give the liquid crystal its properties : These sections are intended to provide the reader with a brief overview of the effects as well as an overview of the literature. The final section on the fabrication of the device by

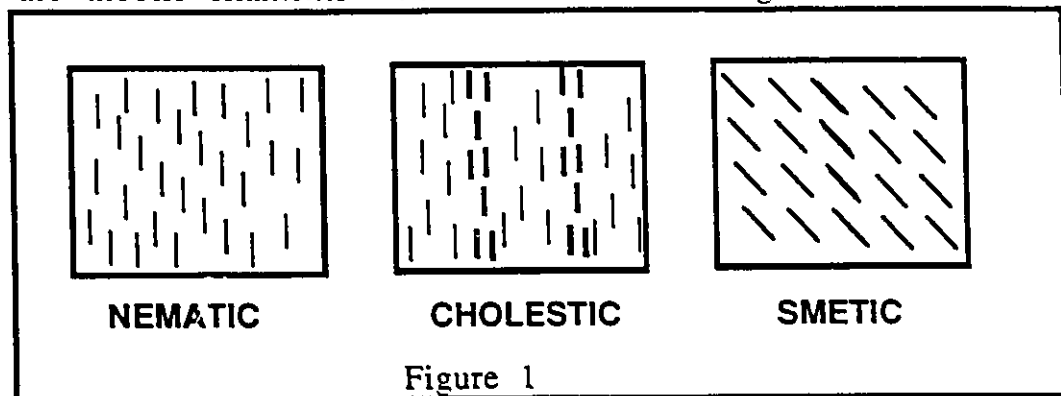
postprocessing of a 3μ Northern Telecom integrated circuit. This section is taken from notes I took while fabricating the device and is very practically oriented in nature. Its purpose is to provide details which are not found in literature in the fabrication of the device, for subsequent iterations of postprocessing.

PHYSICAL AND CHEMICAL PROPERTIES

Liquid crystals are ordered fluids which consist of elongated organic molecules. They look and pour like ordinary liquids. The liquid crystal phase, thermodynamically speaking, is an intermediate phase, a mesophase, which falls between the crystalline solid phase and isotropic liquid phase. The first-order transitions between these phases are sharp, even though they do not occur at the same temperatures upon heating and cooling. There are different categories of liquid crystals, differentiated by the method of preparation. A first category of liquid crystals prepared by heating are called thermotropic. A second category, prepared by dissolving one chemical component in another, for example soap in water, are called lyotropic. Since lyotropic crystals are rarely used, the following discussion will focus on thermotropic crystals [Sore72].

Figure 1 illustrates the three kinds of mesophase ordering; nematic, cholesteric, and smectic. The first ordering, the nematic phase, the molecules have their long axes parallel and are free to slide past one another. Throughout a volume of the liquid, the

average ordering direction can change smoothly. When seen under a polarizing microscope the nematic texture shows dark threads which are mobile filaments where the molecular alignment is discontinuous.



The next ordering is the cholesteric phase. Like the nematic phase, it has the spontaneous ordering, but with a twist superimposed upon it. It consists of parallel layers, each having a definite preferred direction. However, perpendicular to the long axes, the preferred direction is continuously rotating from one layer to the next, thereby tracing out a helical ordering pattern[Sore72].

The third and final ordering, the smectic mesophase structure is similar to a solid. Each layer of the structure has a constant preferred direction in each layer. Within the smectic phase, three different types have been documented. Furthermore, a given substance can occur in several mesophases, such as nematic and smectic, or cholesteric and smectic. In these instances, the smectic phase always occurs at the lower temperature.

Within a mesomorphous compound, nearly every physical parameter is anisotropic, including the electrical conductivity(σ), the dielectric constant(ϵ), the magnetic susceptibility, the viscosity, the elastic modulus, and the optical indicatrix. The σ and ϵ parameter are expressed as having components parallel to and perpendicular to the major molecular axes. The nematic and smectic phases are equivalent, optically, to a positive uniaxial crystal, with the optic axis coinciding with the preferred long-axis direction. Cholesterics are negative uniaxial and have a host of unusual optical properties.

Liquid crystals consist of molecules. These molecules possess the following characteristics : 1) elongated and rectilinear shape, 2) chemical double bond along the long axis, giving rigidity, 3) easily polarizable chemical groups in the molecule together with strong electric dipole moments, 4) weak dipolar groups at the extremities of the molecule. Between molecules there are dipolar attractive forces and liquid crystallinity depends upon a delicate balance of the lateral and terminal intermolecular attractions. The typical example of the liquid crystal chemical composition is two benzene rings joined by a central group, with substituent groups, such as alkyl or alkoxy groups, at the ends of the aromatic rings.

For display applications, the optical anisotropy is important because it ensures that the optical characteristics of the liquid crystal will change markedly when the molecular order is perturbed by an external stimulus. Of the many influences that can perturb a liquid crystal [McVaSi89], an electric field has been found to be the most

convenient means of controlling optical response for display purposes.

ELECTRO-OPTIC MECHANISMS

A large number of electro-optic effects in liquid crystals has been collected for display applications. There are two basic mechanisms operating in any of these effects. The first such mechanism is the dielectric torques and the second is the conduction-induced torques, both of which act upon the molecular orientation pattern. Depending on the magnitude of the applied field, the frequency of electrical excitation and the magnitude of the liquid's conductivity, one interaction will tend to dominate over the other.

The incidence of dielectric torques in insulating liquid crystal can be easily explained. The liquid crystal when in an external field, becomes polarized. The local dielectric polarization P is the vector sum of the induced dipole moment and the permanent dipole moment. Since the dielectric constant is anisotropic, P will in general be oriented at some angle θ with respect to the applied field E ; hence the torque $PE \cos \theta$ is exerted on the molecular pattern and the initial molecular alignment will be unstable for large E . This dielectric torque is a field effect and is associated with displacement currents; the displacement of bound charges in response to electrical stimulus.

Conduction currents and mass flow induced by the external field are all involved in conducting induced torques. In a conductive liquid crystal, the external field produces transport of free charge carriers (conduction current). When the field is active, positive and negative charge carriers segregate, that is, space charge accumulates. The field exerts a force on the space charge which causes a hydrodynamic flow. The flow gradients, also called shear, exert a torque on the molecular orientation pattern, and this shear-induced torque is the most important component of conduction-induced torque, although there is also a dielectric part because the space charge sets up an internal electric field.

There is a large repertoire of dielectric interaction examples: the realignment of a nematic or smectic, the static distortion of a cholesteric, the transformation of a cholesteric to a nematic, the ordering of a disordered nematic, the creation of an oscillating orientation pattern and the alignment of a disordered nematic-cholesteric mixture. Turbulent instabilities produced in nematics smectics or nematic=cholesteric mixtures represent primary examples of conduction-induced torques [Sore72].

DIELECTRIC EFFECTS

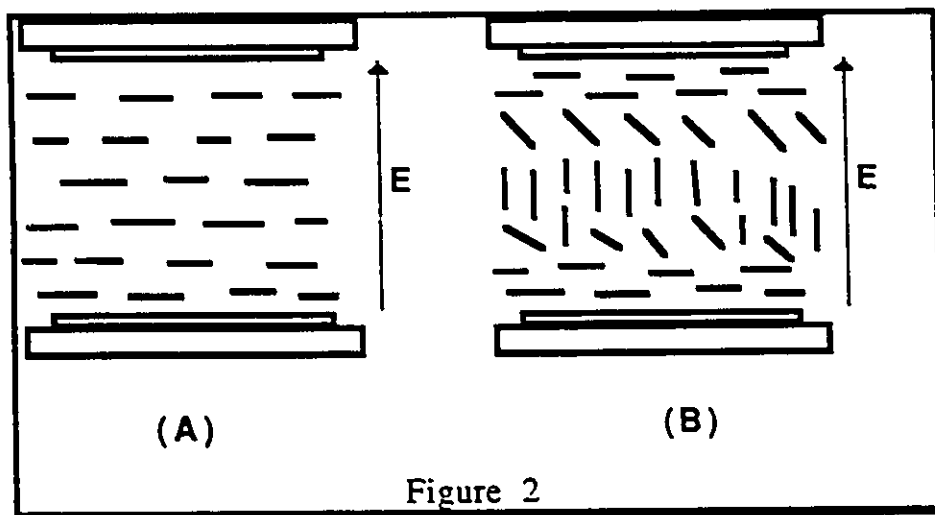
Using the field-induced distortion of a thin film of nematic liquid crystal a light valve can be created. In the following discussion, it will be assumed that the device has the usual parallel plate geometry, with the thin nematic film sandwiched between

transparent electrodes. Two possible methods exist to create the electro-optic device. The first approach has been treated by Goldmacher and Heilmeyer. Here the molecular axes are initially parallel to the walls, the electric field is applied perpendicular to these axes and the nematic has $\epsilon_{||} - \epsilon_{\perp} > 0$ (This molecular alignment would be stable if the dielectric anisotropy were negative). By unidirectional rubbing of the plates with a cotton swab, the alignment parallel to the walls can be produced. The alignment can be perfectly homogeneous or it can consist of randomly oriented swarms whose internal order direction is parallel to the walls.

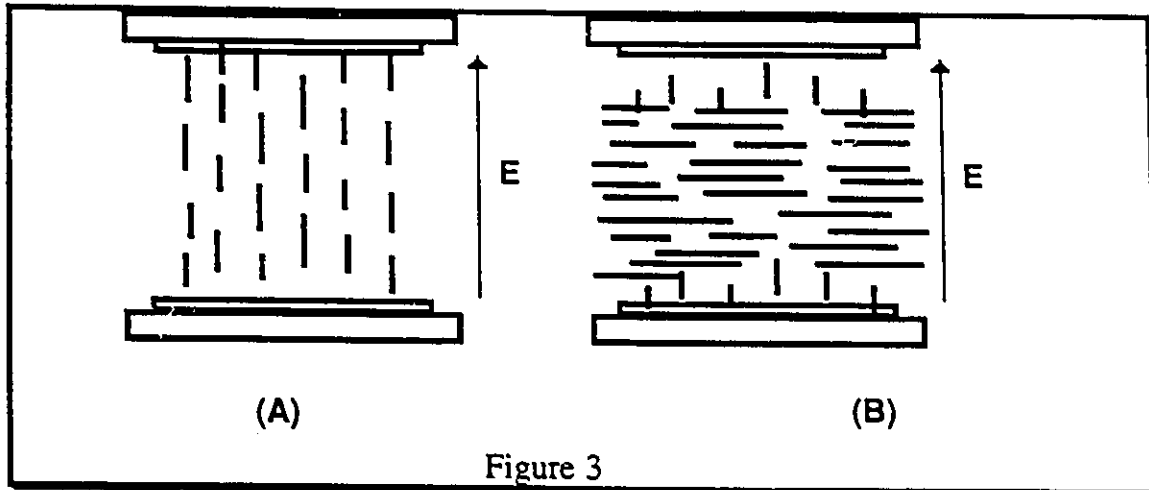
The second approach is that of Soref and Rafuse who have described the approach in which the molecular long axes are originally perpendicular to the electrode substrates, the applied field is parallel to the long axes and the nematic has $\epsilon_{||} - \epsilon_{\perp} < 0$. The forces of attraction between the walls and the ends of the molecules results in the perpendicular, homeotropic ordering. Homeotropy can be promoted by cleaning the transparent electrodes with an acid solution, by using surfactants, and by then rubbing the electrodes with chemically treated swabs [Sore72].

Figures 2 and 3 illustrate the action of the electric field in both cases. In each case, the field is approximately perpendicular to \mathbf{P} . The dielectric torque tends to make \mathbf{P} parallel to \mathbf{E} and hence rotate the molecules. The spontaneous ordering (fig 2a and 3a) is maintained by wall forces and by intermolecular elastic forces. When the field adds enough electrical energy to exceed the deformation

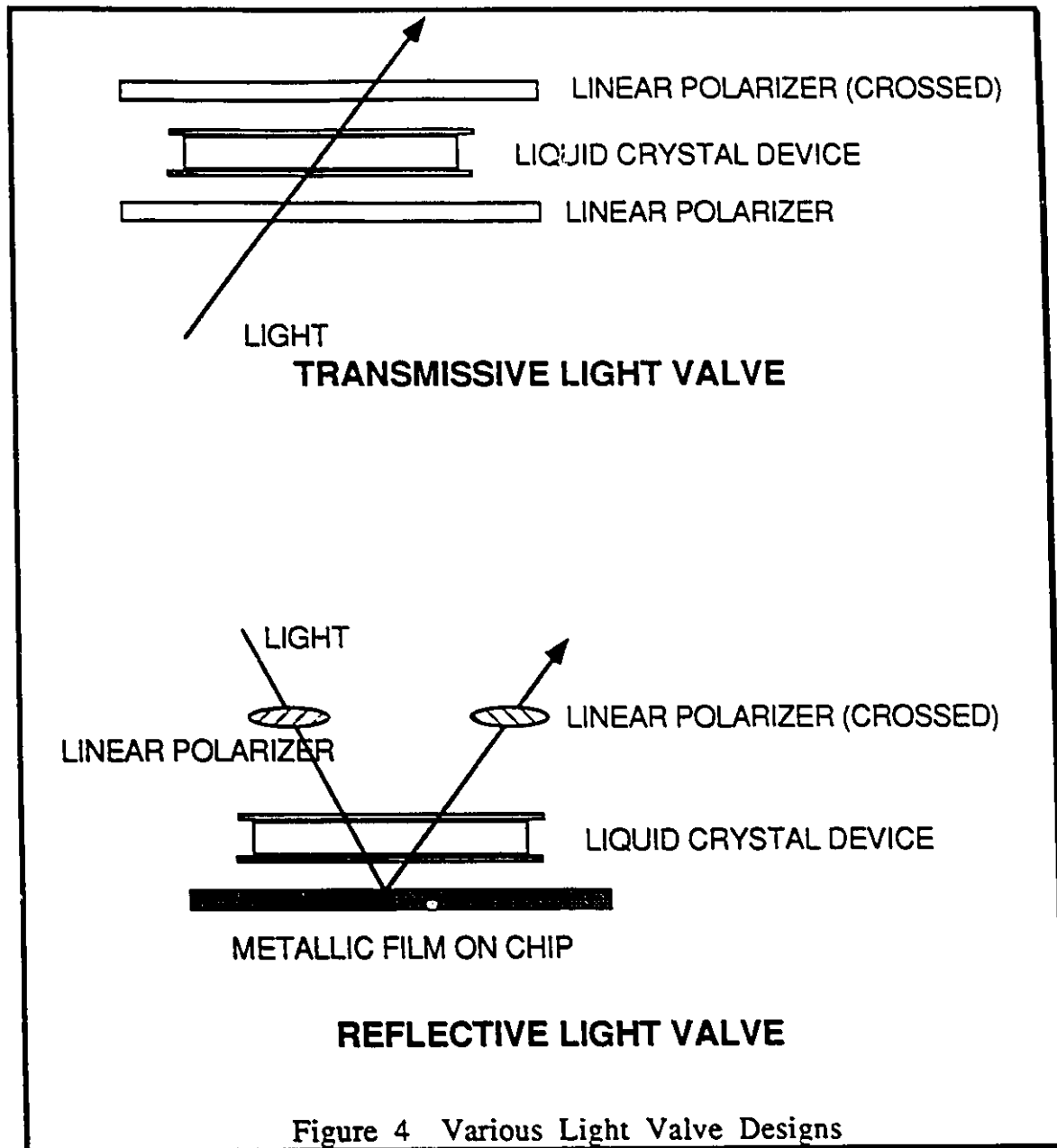
potential of the molecular system, that is, when the threshold is reached, the initial ordering becomes distorted. The molecules realign co-operatively. The change in the molecular pattern at threshold has come to be known as the 'Freedericksz transition'. In figure 2b most of the molecules rotate into parallelism with E, except the molecules in the two surface layers that remain mostly parallel to the walls. In figure 3b, molecules in the central part of the film rotate until they make a large angle (approximately 90 degrees) with the field, although the boundary layers retain their perpendicularity with the walls. The reordering threshold is quite sharp, and is, therefore, useful for display applications.



In figure 3, the opposite sequence of events take place. The birefringence for normal incidence is large below threshold, and diminishes greatly above threshold as the optic axis of the nematic film rotates into near co-incidence with the light propagation.



Consider now the optical effects associated with the molecular reordering. Below the threshold in Figure 36, the birefringence of the nematic layer is very small for light propagating normal to the electrodes (along the optic axis.). The birefringences increases dramatically above threshold, often reaching $\Delta n = 0.6$, which is far larger the the birefringences induced in solid-state materials by external electric fields. In Figure 2, the opposite sequence of events take place. The birefringence for normal incidence is large below threshold, and diminishes greatly above threshold as the optic axis of the nematic film rotates into coincidence with the light propogation direction [Sore72].

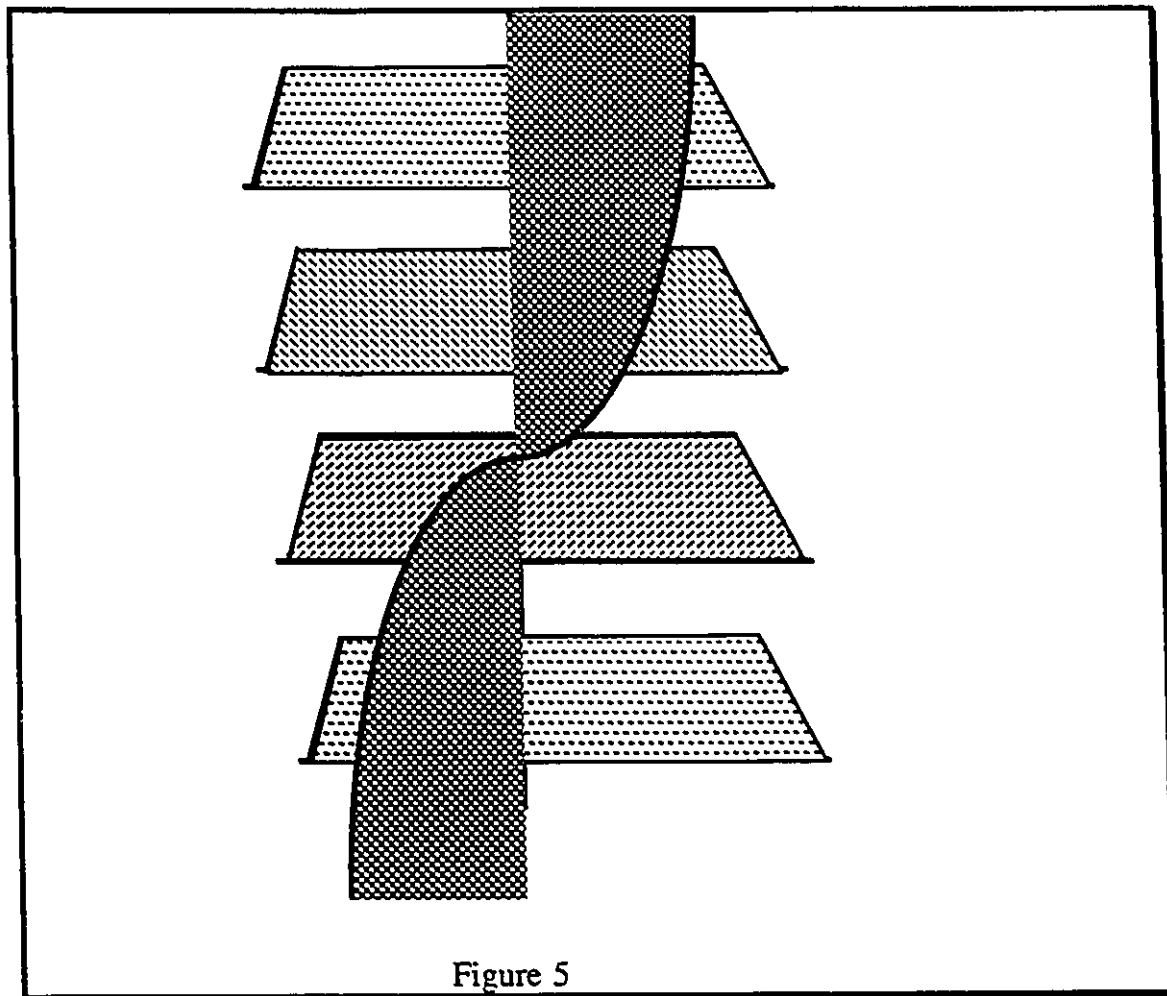


Now that the different effects acting on the liquid crystals in the presence of an electric field have been discussed, how is the electrically controlled birefringence used in light valve displays. Two possible methods are presented in Figure 4. In the transmissive method, where both electrodes are transparent, the nematic device is

placed between crossed linear optical polarizers. The device below threshold is considered in the "field-off" state and the device above the threshold is said to be in the "field-on" state. For the device in Figure 35, light will be transmitted through the crossed polarizer set up in the field-off state because the light beam will become elliptically polarized upon passing through the birefringence nematic film. Light transmission drops greatly in the field on state because of the films near zero birefringence. Therefore, the Figure 4a light valve is open in the field off state and closed in the field on state for the Figure 35 device. By a similar line of reasoning for the Figure 4 device, the Figure 4A valve is closed in the field-off state and open in the field on state. The reflective device employs a circular polarizer and a mirror, or a reflective back-electrode. The circular polarizer consists of a linear optical polarizer and an optical quarterwave plate. The circular polarizer can be replaced by crossed linear polarizers [McVaSi89].

An unusual kind of valve was used in the VLSI implementation of a liquid crystal device. This implementation used a molecular ordering like Figure 2 with $\epsilon_{||} - \epsilon_{\perp} > 0$. except that the top and bottom plates were rubbed in orthogonal directions. This caused the ordering direction of the material to twist θ degrees about \mathbf{E} in going from one electrode to the other as shown in Figure 5. Moreover, it has been shown that the pattern would "untwist" above a threshold field of 2.5Vdc. This property will be used in the construction of our light valve, with the incoming optical electric vector oriented parallel to one of the rubbing directions. For the material used and a search of

current literature in the field , resulted in using a 45 degree twist between electrodes for the post-processed device.



CONDUCTION-INDUCED EFFECTS

The wide-angle light scattering that accompanies electrohydrodynamic turbulence in liquid crystals is called dynamic scattering. The turbulent liquid crystal contains moving birefringent regions several microns in size, and the diffuse optical scattering

comes from the irregular refractive index gradients that are present. As the voltage is raised across the nematic layer, the conductive nematics go through two hydrodynamic instability thresholds. The first, domains appear and then, at higher voltages, turbulence sets in. The domains are also called Williams stripes, cellular flows, or vortex cells, and are above the scope of this discussion; these phenomena are discussed in [Sore72].

The necessary and sufficient conditions for nematic to exhibit domains and turbulence include factors such as the conductivity must be greater than a minimum value, around 10^{-11} (Ohm cm)⁻¹. Originally, it was thought that $\epsilon_{||} < \epsilon_{\perp}$ is also a necessary condition, but recently it has been shown this factor is oversimplified because they obtained domains in certain liquid crystals which has $\epsilon_{||} > \epsilon_{\perp}$.

Dc or ac excitation can result in electrohydrodynamic instabilities, and it is believed that the dc and ac instability mechanisms are different. The mechanism is thought to be dominant at dc. In the ac regime, the model of Carr and Helfrich gives good agreement with the experimental data and the Carr-Helfrich mechanism also contributes slightly at dc.

Felici's model assumes that current is injected into the liquid crystal. Controversy still exists as to how the charge carriers are injected, but present evidence favors the idea of electron injection at the cathode. Injected electrons can attach themselves to neutral molecules and be carried as negative ions. There is a theory that

impurity ions in these liquid crystal migrate to an electrode without discharging thus forming an electrical double-layer at the electrode-electrolyte interface. This promotes injection by lowering the potential barrier between the liquid and electrode.

A net space charge Q per unit volume builds up, which couples to the electric field E with a body force QE . From momentum conservation, the resulting ion motion is counterbalanced by the flow of neutral liquid crystal molecules. The fluid speeds up as the applied voltage is increased, and at a critical voltage, a structural instability in the flow pattern is reached; domains are formed. Conservation of flow requires that the fluid streamlines turn around at the electrodes, thereby setting up a spatial pattern of rotary flows with opposite rotation in adjacent cells. It is interesting to note that the Felici mechanism does not rely upon the liquid crystal anisotropy and that electrohydrodynamic flows are observed in the isotropic phase, as predicted by the model [Sore72].

OPTICAL STORAGE AND OTHER CONDUCTION-INDUCED EFFECTS

The phenomenon of optical storage occurs in mixed nematic-cholesteric systems that are roughly 90% nematic by weight. Addition of cholesteric material may increase the nematic's resistivity, but more significantly the mixture acquires many of the properties of a cholesteric liquid.

When a mixture is dc or ac excited, below the dielectric relaxation frequency, it is at this point when storage comes about. As in the nematic case, mobile ions are available for conduction whether coming from injection, impurities, or dissociation. The field produces transport of these ions through the mixture and, as above, the associated shear leads to hydrodynamic instabilities. There are two instabilities which are observed. The first occurs at low voltage. The microscopic appearance of the liquid crystal layer changes from loosely threaded appearance to a set of irregular stripes, similar in appearance to a fingerprint. These striations are spaced approximately two microns apart, unlike the parallel rectangular nematic domains, whose width is about twice the layer thickness. With increasing voltage, the molecular pattern becomes more irregular and agitated, giving way to turbulence at about three times the domain voltage. The turbulence disrupts the molecular texture, and focal conic cholesteric regions begin to nucleate where the disruption is severe. After the electric field is switched off, the focal-conic areas remain for days or weeks, in the field state, and this texture constitutes the diffusion light scattering, memory state[Sore72].

Electrical erasure of the stored scattering state is dielectric interactions. Since the mixture has negative dielectric anisotropy, an applied high frequency field will align the molecular long axes parallel to the electrode which returns the mixture to the Grandjean plane texture. If the helix pitch is properly chosen, the characteristic

reflection colors will fall outside the visible spectrum and the liquid crystal layer will be transparent.

Electrically induced turbulence, storage, and erasure have also been found in smectic-phase materials, either as the polygon or simple fan texture, but not in the mass texture. The mechanisms involved are believed to be quite similar to those discussed above. Some of these phenomenal smectics may turn out to be useful in displays, but further research is needed to determine whether this is so.

DEVELOPMENT OF A SPATIAL LIGHT MODULATOR

Liquid crystal spatial light modulators (SLM) and displays perform a similar function in two distinct areas of application, namely, the modulation of light for optical information processing or for displaying visual patterns. The device specifications for a display and SLM are, however, slightly different. A display usually operates with noncoherent polychromatic/white light in an environment where it may be subject to temperature changes (10°C to 30°C) and often where a wide viewing angle is required. A SLM may be operated with monochromatic light at nearly normal incidence in a temperature stabilized enclosure. Also in the case of a SLM, higher optical contrast between ON and OFF states and faster switching speeds are usually desirable. For coherent optical processing applications both the amplitude and phase modulations of the light must be specified precisely pixel by pixel, even for binary

modulators. This imposes the constraint that the light transmitting regions, both passive and active, should have an optically uniform response across the device, and in addition in the case of reflective SLM the reflecting surfaces should be optically flat [McVaSi89].

Various types of optically and electronically addressed spatial light modulators have been presented over the last decade. The SLM that is to be discussed here is similar to some of these in that it contains electronically addressed pixelated devices, but from what I can find, this is the first to be implemented in Canada using the Northern Telecom fabrication process. Electronically addressed SLMs may be classified as static memory or dynamically addressed devices depending on whether a static memory cell is incorporated into each pixel of the array. Static memory arrays are expected to provide more stable drive conditions for the liquid crystal layer and, therefore, better optical performance but at the expense of increased complexity and power dissipation. However, this is offset by the simpler addressing protocol, since there is no need to map continually the pattern to be displayed from an external frame store onto the device just to maintain the display as in the case of dynamically refreshed SLMs. Static memory devices have their own built in frame stores, and need only to be addressed when the pattern requires updating.

The silicon backplanes developed originally by the display industry usually consist of a matrix array with a single transistor per pixel to provide charge storage. They are addressed a row at a

time, and the minimum frame scan time is set by the number of rows and the time to charge a pixel. For certain liquid crystal configuration such as high speed ferroelectric liquid crystal materials the high spontaneous polarization may preclude the use of such dynamically refreshed backplanes, in which case a static memory cell at each pixel is essential to drive the new faster materials [McVaSi89].

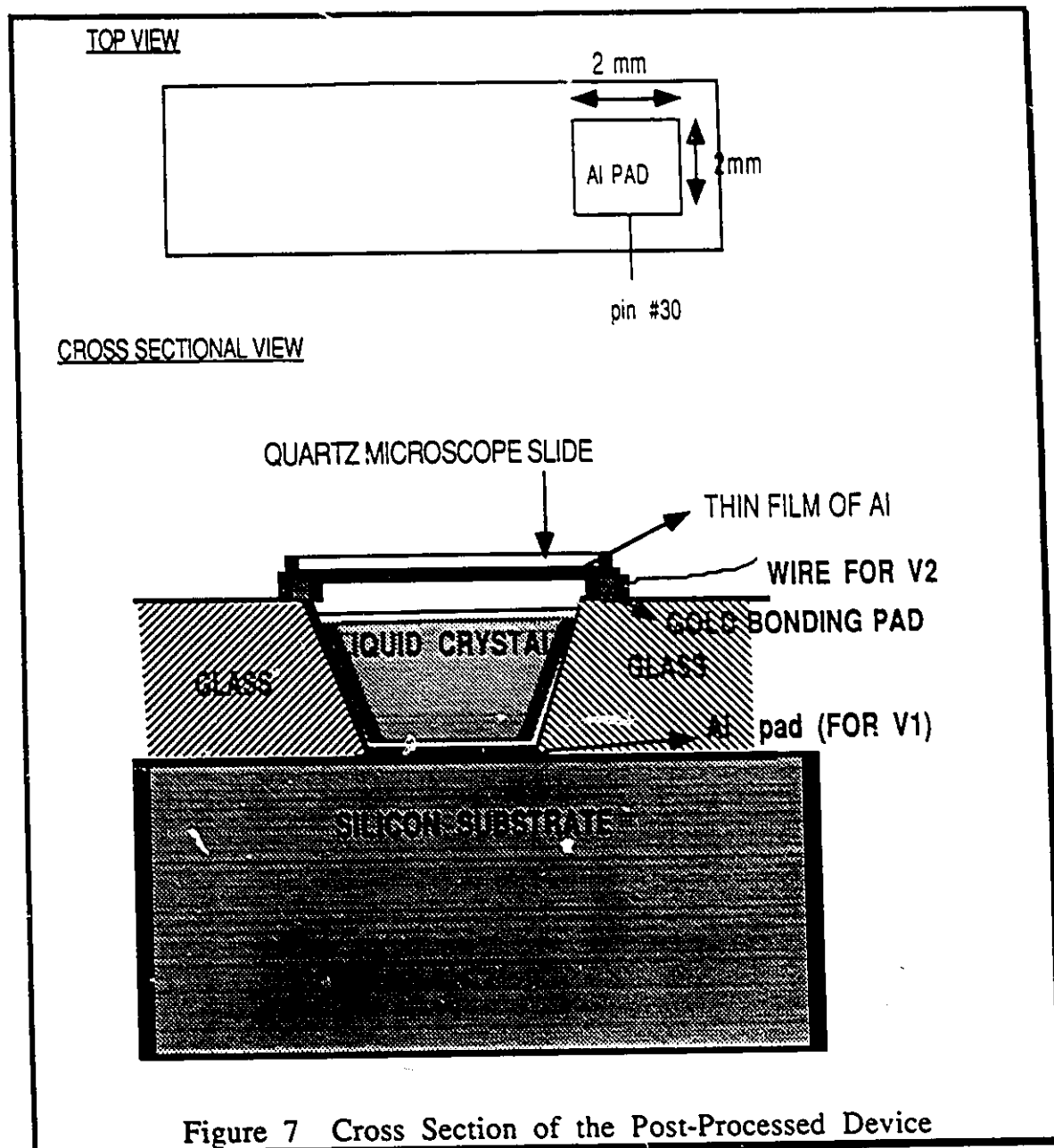
STRUCTURE OF THE DEVICE

The chip WRONN contains the basic foundation for a spatial light modulator using standard Northern Telecom layers. Silicon wafers have been processed using the 3micron CMOS process being processed at Northern Telecom. The display pixel consists of a metal pad 2mm by 2mm formed by evaporating a metal 1 onto the silicon wafer. The metal pad not only acts as an optically flat mirror, but also as an electrode for transmitting the voltage drive signals from the output of the pixel circuitry to the element of the overlying liquid crystal layer immediately above it. Adjacent pixels in future submissions should be placed no closer than 100 microns away to prevent optical interference patterns. A mask layout to provide a "well" on the surface of the chip is implemented using the GLASS layer.

The silicon chip is mounted in a standard 128 pin package with gold wires connecting the metal conducting pad of the pixel. A quartz plate with an area of 5.0mm by 5.0mm is fitted just clear of the

bonding pads to enclose a layer of liquid crystal, BDH type E7 (helical structure shown in Figure 37, over the active area of the SLM pixel. The inner surface of the quartz cover has a coating of indium-tin-oxide (ITO), which forms a transparent electrically conducting film of sheet resistance of 600 ohms/squ, and an aluminum layer evaporated around the edges and sides of the quartz cover slip provides an electrical contact between the ITO electrode and the external circuitry [McSiVa89].

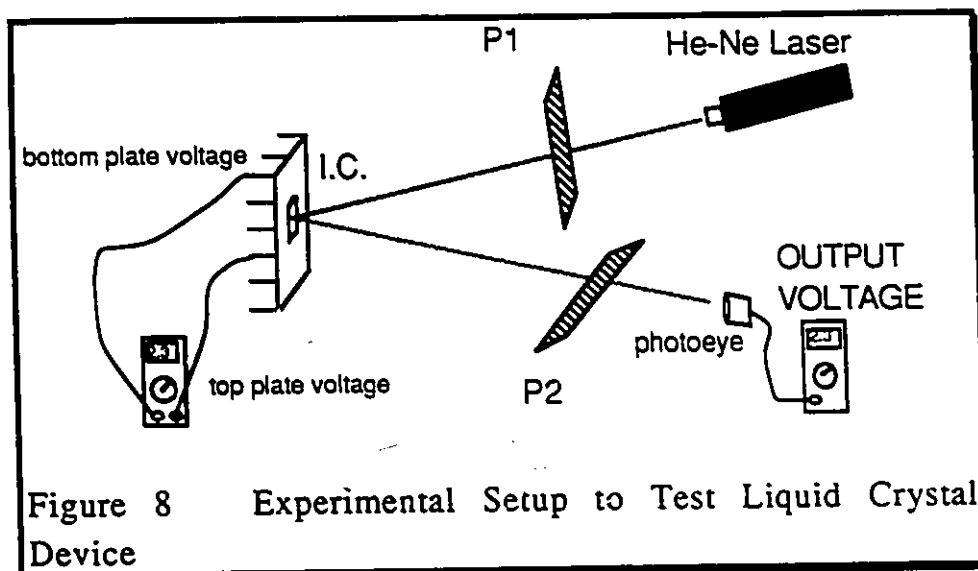
Magnesium Fluoride or SiO, obliquely evaporated at 30 degrees onto the surfaces and deposited to a thickness of .12um, covers both the ITO coating and the surface of the silicon chip. The interactions between the SiO (or MgFl) material and the liquid crystal molecules produce a strong alignment of the liquid crystal at the boundaries of the layer. In other words, this layer untwists the twisted nematic structure. The cell is assembled so that there is an angle of 45 degrees between the alignment directions at the glass and silicon boundaries. Therefore, with no voltage applied, the liquid crystal layer exhibits a helical structure with a twist of 45 degrees across the cell. The width of the layer is determined by the thickness, 12um, of the mylar spacer around the edges of the active area of the SLM array. Epoxy resin is used to hold the assembly together(Figure 7shown below).

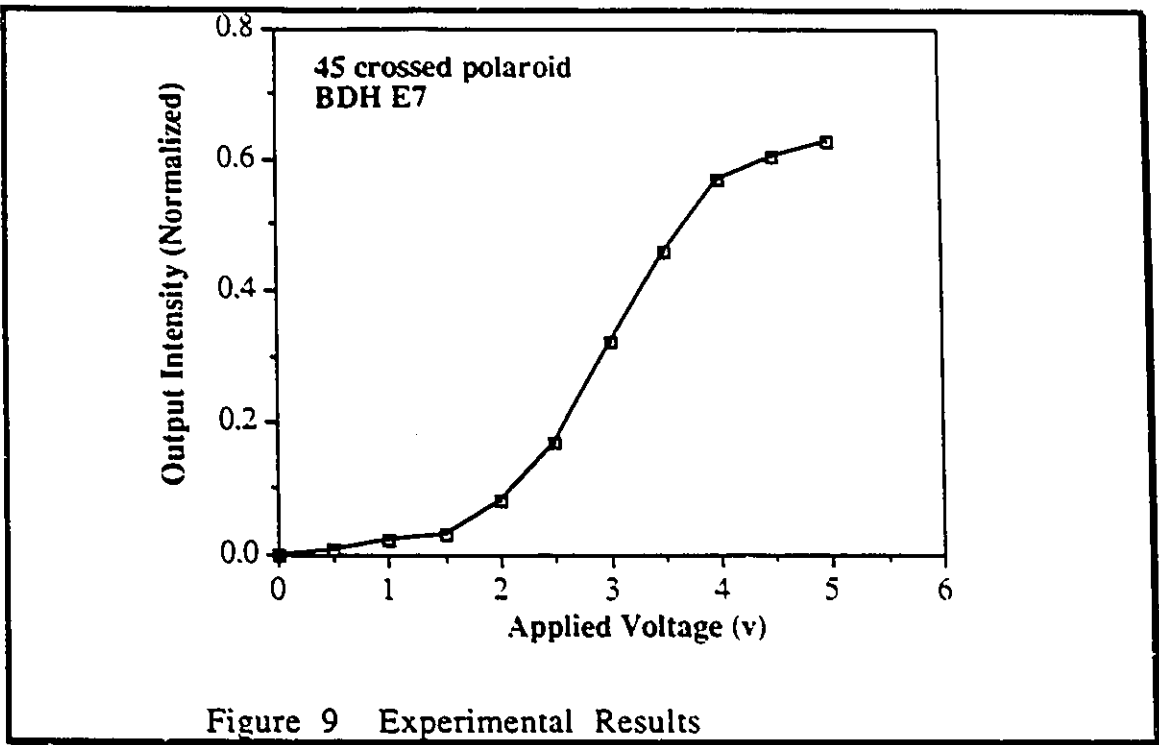


LIGHT MODULATING PROCESS

When plane polarized light is incidentally normal on the SLM, with the plane of polarization parallel to the alignment direction at the glass boundary, the helical structure in the liquid crystal layer

guides the polarization vector very efficiently as the light propagates, so maintaining the light as a plane polarized wave but in a rotated plane. With no voltage applied across the liquid crystal layer the polarization guiding property of the structure ensures that on reflection from the SLM the plane of polarization of the reflected wave coincides with that of the incident wave. By including a crossed Polaroid it is arranged that the output beam from the system has almost zero intensity as shown in Figure 8. When an oscillating voltage signal is applied across the liquid crystal layer with sufficiently high amplitude to modify the helical structure so that complete polarization guiding no longer occurs, light is coupled between the two orthogonal polarization states. The light reflected from the SLM is not in general elliptically polarized, and light is transmitted through the crossed analyzer in this case- the ON case (Figure 9).





OVERVIEW OF THE DEVELOPED DEVICE

Clearly then, a light has been developed using an aluminized silicon wafer as the reflecting back electrode and an ITO coated glass plate as the transparent front electrode. This device is to be used as an output device for the VLSI implementation of a neural network. The cells are assembled with a 45 degree twisted nematic arrangement. The post processing has taken place at the Liquid Crystal Institute located at Kent State University. I would especially like to thank Dr. Jack Kelly for use of his clean room and chemicals, as well as his patience and kindness.

APPENDIX III

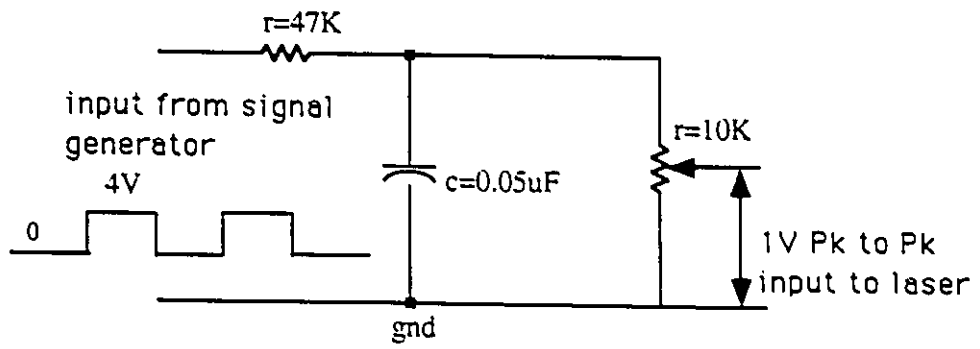
SPIICE DECK AND MODEL CARD FOR SIMULATING PHOTORECEPTOR

```
.model npn nmc >(bf=250 br=1 is=1e-16 vto=0.7 kp=40e-6  
+ phi=0.6 lambda=0.01 re=0.5 rc=120 gamma=1.1 va=100  
+ rd=40 rs=40 pb=0.7 cgso=3e-10 cgdo=3e-10 cgbo=5e-10 rsh=25  
+ cj=4.4e-4 mj=0.5 cjsw=4e-10 mjsw=0.3 js=1e-5 tox=5e-8  
+ nsub=1.7e16 tpg=1 xj=6e-7 ld=3.5e-7 uo=775 vmax=1e5 level=1)
```

```
bjt1 Thursday, March 13, 1990 3:48 PM
.include cmos3d.lib
bjt1 vout base vdd npn
mn1 vout vout v1 0 ntran w=5u l=5u
mn2 v1 v1 v2 0 ntran w=5u l=5u
mn3 v2 v2 v3 0 ntran w=5u l=5u
mn4 v3 v3 0 0 ntran w=5u l=5u
vdd vdd 0 dc 5.0
```

```
.end
```

APPENDIX IV
CIRCUIT FOR TRIGGERING LASER



Vita Auctoris

Naveen Bewtra was born on August 5, 1965 in New Delhi, India. He completed his high school education at Assumption College in Windsor in 1984. In 1988 he graduated from the University of Windsor with a Bachelor of Applied Science in Electrical Engineering. In June 1988, he was employed at Northern Telecom Canada as a Test Engineer for the DMS Supernode. After being awarded an educational leave of absence from Northern Telecom, he completed his Masters of Applied Science in Electrical Engineering, also at the University of Windsor. Mr. Bewtra is an IEEE member.