

# **A CMOS Foveated Image Sensor**

Robert Wodnicki

B.Eng, (McGill University), 1992

Department of Electrical Engineering  
McGill University  
Montréal  
January 31, 1996

A Thesis submitted to the Faculty of Graduate Studies and Research  
in partial fulfillment of the requirements of the degree of  
Master of Engineering

© Robert Wodnicki, 1996

8.11	Temporal and spatial noise in the periphery array . . . . .	131
8.12	Example images from the periphery array . . . . .	132
8.13	Noise in the prototype system . . . . .	135
8.14	Black and white bars imaged by the fovea . . . . .	137
8.15	Spatial noise in the fovea array . . . . .	139
8.16	Sample images for the fovea array . . . . .	141
8.17	Charge spreading in overilluminated wristwatch image . . . . .	143
A.1	Experimental optical setup . . . . .	155
A.2	Image of light probe spot . . . . .	156
B.1	Photomicrographs of microsurgery on the periphery array . . . . .	158
B.2	Schematic of surgery on the fovea array . . . . .	161
B.3	Photomicrographs of microsurgery on the fovea array . . . . .	162

## List of Tables

8.1	Response of test chip photodiodes . . . . .	117
8.2	Factors limiting the speed of the periphery array . . . . .	133
8.3	Power consumption of periphery array . . . . .	136
8.4	Factors limiting the speed of the fovea array . . . . .	145
8.5	Power consumption of the fovea array . . . . .	146

# Chapter 1

## Introduction

### 1.1 Motivation

Future mobile robots are expected to demonstrate a greater awareness of their environment through the use of *computer vision* algorithms presently under development. With the help of machine vision sensors, mobile robots would be able to detect looming objects well ahead of a collision. Such mobile robots would ‘see’ far greater distances than is possible with conventional range imaging. They would be able to map out their immediate surroundings passively without the need for expending energy to drive laser or SONAR range finders. With the help of higher-level algorithms, mobile robots would begin to recognize objects in the world around them and use this information to direct their own actions. The potential exists for the application of advanced prototypes of these machines in dangerous circumstances such as space exploration and maintenance of deep sea pipe-lines.

Before such ‘seeing’ mobile robots can become viable, they must become autonomous. Current systems function well in the laboratory where they can be tethered to powerful host computers either through the use of cables or with radio data links. Navigating in hostile environments however, often requires fast reaction to external stimulus. In some cases — such as the exploration of Mars — the time it takes for a signal to be transmitted from the robot to the host computer may be longer than the reaction time for survival. Therefore, successful mobile robots would be required to interpret sensor data using on-board computing resources alone. They would also have to be extremely fuel efficient to survive for long periods of time on their own



power. These requirements dictate the use of high performance computer vision systems which are small, light-weight and consume very little energy.

## 1.2 Efficient vision sensors

The design of future vision sensors must take into account restrictions on system size and efficiency imposed by autonomous operation. These requirements can be met through a drastic reduction in processing overhead, and by system implementations which capitalize on the use of a highly integrated low-power imaging technology.

A proposed solution to the problem of excess data processing overhead is the use of a space-variant image sensor. This technique concentrates the attention of the vision system on interest features in a given scene. These features are assumed to contain the most important information to be sensed and are therefore resolved at the maximum resolution available to the system. Areas in the perceived scene other than these interest features are assumed to contain less important data, and are therefore sensed at a lower resolution. This scheme significantly reduces the amount of image data which must be processed by subsequent higher level vision algorithms, while at the same time preserving the overall field of view of the system. It therefore results in a significant savings in power consumption and system mass.

The above described system may be implemented using standard image processing hardware using a CCD camera for image capture however further savings could be realized through the use of a low-power, highly integrated image sensing technology. Bringing together image sensing and image processing functions in one device can serve to reduce excess power dissipation due to communication between previously separate units such as CCD camera, frame-grabber, and computer. The elimination of redundant power supplies, wiring, and interface circuitry in such a system can lead to a significant reduction in system mass and result in a further reduction in power consumption of the mobile robot. Finally, consolidating sensing and processing functions in one condensed system allows for parallel architectures which result in significant increases in computing power and system through-put.

Therefore, the synergy of space-variant sensing and new highly integrated imaging technology is expected to yield considerable amelioration in the performance of machine vision systems for autonomous mobile robots. This thesis investigates the potential of these new vision sensors. The algorithm for data reduction is called foveation and is motivated by studies of data reduction present in the human visual pathway. The technology for realizing highly integrated, low-power vision sensors is the emergent CMOS imaging paradigm.

### 1.3 Thesis organization

The thesis begins with a study of the foveated mapping in Chapter 2. A model based on biological evidence and the restrictions of a physical implementation is developed, and a discussion of the properties of the algorithm is presented. In addition a discussion of the use of foveation in mobile robot applications is presented.

Chapter 3 examines silicon phototransducers in an effort to provide adequate background for the study of CMOS image sensors later on. Some basic concepts in optoelectronics such as responsivity and photogeneration in silicon photodiodes are discussed. This is followed by a survey of the various phototransducers available for implementation in standard CMOS technology.

Chapter 4 continues the introduction to CMOS imagers with a brief chronology of the development of solid-state imaging. It begins with a discussion of the operation of the Charge Coupled device (CCD) image sensor, and follows with an analysis of the problems with this technology. The remainder of the chapter is devoted to a survey of present research into the use of standard CMOS processes for the realization of image sensors for computer vision.

In Chapter 5 the simplest CMOS imager, the Passive Pixel Sensor (PPS) is introduced. The basic architecture of this image sensor is presented and its three main components are then examined in detail. The theoretical discussion is followed by the simulation of an  $8 \times 8$  PPS imager using the Spice [1] circuit simulation software package.

Some factors limiting the expected ideal performance of CMOS imagers are examined in Chapter 6. The effects of clock feed-through and MOSFET threshold voltage  $v_T$  are investigated. The two types of noise which affect CMOS imagers—temporal, and spatial noise—are examined. Useful metrics for quantifying noise are then presented.

The design of the prototype CMOS foveated image sensor is discussed in Chapter 7. The foveated mapping is implemented by realizing separate PPS arrays for the fovea and periphery. Various engineering trade-offs undertaken in the design process are discussed with the help of photomicrographs of the fabricated device. The design of the high resolution fovea array is examined, and an account of the development of C language software (**PGEN**) for automation of the layout of the periphery array is presented.

Chapter 8 provides a comprehensive discussion of the test of the CMOS foveated image sensor. It begins by examining electrical and opto-mechanical test apparatus developed to properly evaluate the prototype sensor. Calibration procedures and sampling and processing of image data are discussed. Sample images taken with the fabricated prototype are then presented. Examples of these are shown in Fig. 1.1. These and other images are examined in detail to evaluate the quality and utility of the prototype sensor. Performance measures such as maximum frame rate, power dissipation, and dynamic range are presented to complete the analysis.

The thesis closes in Chapter 9 with a summary and discussion of key issues. Suggestions for future work are presented in an effort to stimulate further research in the use of standard CMOS for the implementation of foveated imagers.

## 1.4 Contributions

This thesis presents a study of a novel CMOS foveated image sensor for computer vision. Many of the ideas expressed are well known and have been reviewed in order to provide sufficient background for the analysis of the sensor. There are however three main concepts which, it is felt, have not appeared elsewhere before, and the

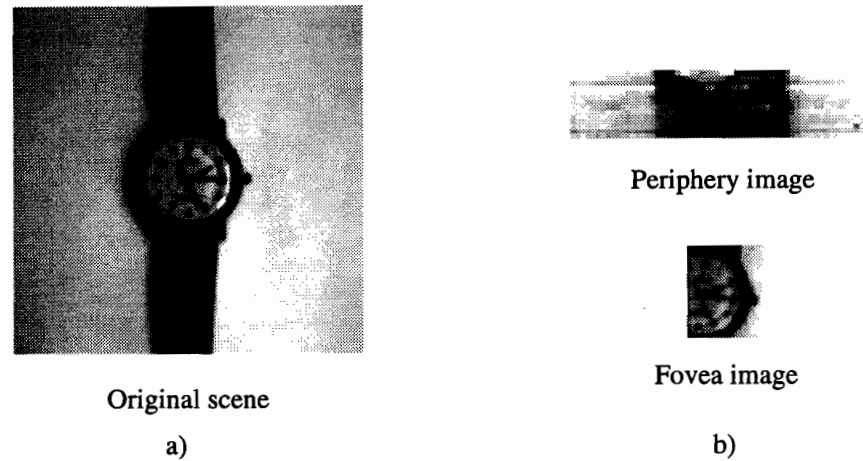


Figure 1.1: Foveation with the prototype sensor. a) The original scene is a wristwatch against a white background. It contains regions of fine detail (such as the numbers on the watch face), as well as regions with less detail (the watch band). b) The outputs of the sensor are in the form of separate *fovea* and *periphery* images. In typical use, the attention of the sensor would be focused on the center of the watch face. The fovea output is a high resolution image of the detailed part of the original scene, whereas the periphery output is a log-polar representation of surrounding features.

author wishes to claim these as original.

To date a number of foveated sensors have been proposed in the literature, however the present work is believed to be the first fully documented sensor to be implemented in a standard CMOS process [2]. The author claims originality for the express use of CMOS imaging technology to implement a foveated sensor for mobile robots in order to capitalize on savings in power, system cost and system size.

In order to properly realize the foveated mapping in a standard CMOS technology a new model of foveation was developed. The author claims originality for the *hybrid* model of foveation discussed in Chapter 2 including the derivation of the consequences of non-overlapping receptive fields.

Finally, in order to implement the hybrid model in standard CMOS, the problem of the layout of the periphery array with non-Manhattan style geometry had to be solved. The author claims originality for the complete design of the CMOS foveated image sensor, including the novel use of C language software (**PGEN**) and modular unit cells for layout of the periphery array.

## 1.5 Reader's guide

To assist readers seeking specific information from this thesis, a short guide has been compiled with suggestions for effective use of the material.

- For a brief introduction to the foveated mapping, readers should study Sections 2.1-2.2 and 2.6.
- Readers interested in a more detailed understanding should read all of Chapter 2. They can then progress directly to Chapter 7 to gain insight into the design process for migrating the algorithm to silicon. Chapter 8 provides sample images from the fabricated prototype in Section 8.4.2 and Section 8.5.2.
- For a brief introduction to CMOS imagers, readers should study Sections 4.5-4.6.
- Readers interested in a more detailed discussion of this technology should first review the material presented in Chapter 3 and Chapter 4. Chapter 5 should then be read in conjunction with Chapter 6 to obtain an appreciation of some of the merits and limitations of a basic CMOS imager architecture.
- To fully understand the design of the sensor, it is useful to read Chapter 7 in conjunction with Chapter 5.
- To fully understand the test of the sensor, it is necessary to read Chapter 8 in conjunction with Chapter 6.
- Readers interested in the differences between hardware and software implementations of the foveated mapping should read this thesis in conjunction with accounts of similar work. For a comparison with a solid-state CCD implementation, the work of Kreider *et al* [3] is very useful. For a comparison with a real-time software implementation, the work of Bolduc and Levine[4] provides an excellent discussion of a prototype system.

The thesis may of course be read straight through from beginning to end in which case it functions as an involved discussion of the methodology used to design and test the prototype CMOS foveated image sensor. This discussion begins in the next chapter with an introduction to the foveated mapping.

## Chapter 2

# The Foveated Mapping

### 2.1 Introduction

Foveation evolved out of the results of experiments conducted on primates and cats in the middle part of this century. Most authors in this field quote work by Daniel and Whitteridge [5], Talbot and Marshall [6] and Hubel and Wiesel [7]. These researchers investigated the relation between the retinal visual field and the cortical visual field using point mapping experimental techniques. This neurophysiological data suggests that, contrary to common experience, mammalian vision is not a uniform one to one mapping of coordinates.

Instead, the retina is composed of a central high resolution sensing area, known as the *fovea*, and a surrounding low resolution area known as the *periphery*. Moreover the actual mapping of these two regions onto the visual cortex of the individual is nonlinear, and may be described by a conformal mapping under a complex logarithm function. This effect is known as *foveation* and is responsible for a considerable reduction of data complexity in the human visual pathway.

The consequences of foveation are far reaching and may even be associated with image understanding. The function is found to obtain a high degree of image data compression while still preserving important features of the visual scene. It is also ‘form-invariant’ [8], displaying the useful feature of mapping magnifications in the retinal domain to shifts in the cortical domain, and rotations in the retinal domain to translations in the cortical domain. Consequently it is ideally suited to such high

level visual tasks as character recognition and optical flow tracking [9].

In this chapter, the foveated mapping is introduced and the motivation for its use is examined. Section 2.2 reviews the biological basis for the mapping. Section 2.3 shows how this data has been elaborated to mathematical models, and provides the background for the model used later on in this thesis which is developed in Section 2.4. Implementations of these models in software and hardware are examined in Section 2.5. The chapter concludes with a brief discussion of foveation as it relates to robotic systems in Section 2.6.

## 2.2 Biological evidence

It is known that the spatial distribution of receptive fields in the human retina is non-uniform. The highest concentration of receptive fields appears in a region occupying approximately 5.2 degrees of the entire 204 degrees of visual field [10]. This region, known as the fovea, is responsible for the high resolution component of our vision. Outside of this area, visual acuity decays rapidly as the visual world is segmented into receptive fields whose size grows approximately linearly with eccentricity [11]. This second region is known as the periphery.

In the fovea each photoreceptor is mapped to only one bipolar cell thus defining a receptive field. In the periphery however many individual photoreceptors are grouped together, all sharing one bipolar cell. In this way the response of a large group of photoreceptors is combined into a single output representing a spatially distributed receptive field.

The advantage of such an arrangement is a tremendous reduction in data handling overhead without sacrificing the size of the overall visual field. Instead of hundreds of millions of fibers in the optic nerve, there are only of the order of a million. The amount of brain tissue necessary to process this information is likewise reduced. Therefore it would appear that considerable economy in computer processing might be realised if this system model were to be borrowed from biological vision and adapted to machine vision applications. The first step towards such a system is

an exact mathematical model of the biological data.

## 2.3 Mathematical models

One of the first researchers to propose a mathematical description of the biological data is Schwartz [12] who has made extensive use of the work of Daniel and Whitridge as the foundation for his theoretical development. Their experiments relate retinal coordinates in degrees of visual field to cortical coordinates in millimeter translations along the surface of the primate visual cortex in point matching/probing type experiments. They use the results of these data gathering experiments to create a plot of what they term the “cortical magnification” (in mm/degree) [5]. Schwartz finds this data to fit the following relation,

$$M(r) \approx k/r \quad (2.1)$$

where  $M$  is Cortical Magnification,  $r$  is radial eccentricity, and  $k$  is a constant.

Schwartz postulates that, since cortical magnification is a differential quantity it represents the scalar gradient of some conformal, isotropic mapping. He then goes on to show that the integral of this variation must obey a natural logarithm type law as follows,

$$w = \ln(z) \quad (2.2)$$

He claims that this is a reasonable approximation to the data of Daniel and Whitridge.

The Schwartz model has the form shown in Fig. 2.1. Points in the input image constitute the  $z$  domain. These are mapped to the output image on the  $w$  plane under the conformal mapping given by Eq. (2.2). Fig. 2.1 clearly illustrates how Schwartz’s model also conforms to the data of Talbot and Marshall [6]. Their experiments show that radial lines are mapped to more or less straight lines in the cortex and that exponentially spaced circles in the retinal field are mapped to more or less equally spaced straight lines in the cortex. These features are reproduced faithfully by the  $\ln z$  mapping and are a reflection of the scale and rotation invariant properties



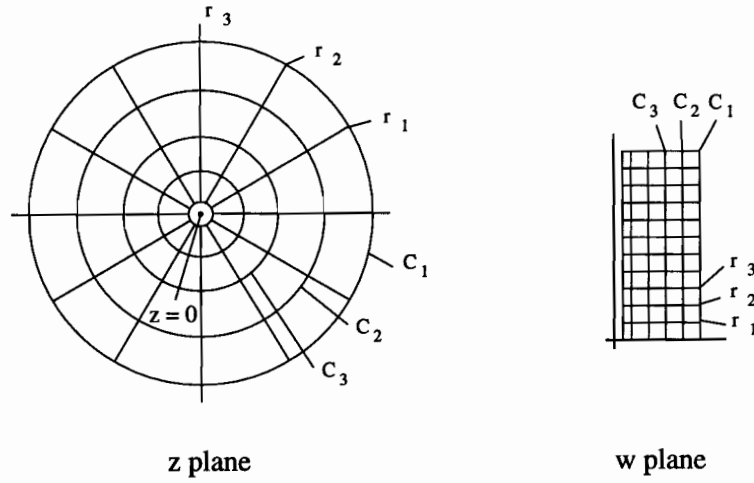


Figure 2.1: The  $\ln z$  mapping. Exponentially spaced circles in the  $z$  plane ( $c_1, c_2, c_3$ ) map to vertical lines in the  $w$  plane. Radial lines in the  $z$  plane ( $r_1, r_2, r_3$ ) map to horizontal lines in the  $w$  plane. The point  $z = 0$  produces a singularity in the  $w$  plane.

inherent in the model.

While it appears that Eq. (2.2) accurately describes the mapping contained in the visual pathway, the presence of a singularity in the  $w$  domain for the point  $z = 0$  is a source of contention. To solve this problem, Schwartz proposes a slight modification to Eq. (2.2),

$$w = \ln(z + a) \quad (2.3)$$

where  $a$  varies between  $1.7^0$  and  $0.3^0$  for the model to fit the biological data [13].

Although Eq. (2.3) solves the problem of the singularity for  $z = 0$ , it does so at the expense of a slightly warped output image. Furthermore, for a hardware implementation, as for the sensor described in Chapter 7, the physical constraint of a minimum pixel size becomes problematic near the origin.

Sandini *et al* address this issue by essentially dividing up the retinal map [11]. They find that the periphery may be generated by sampling the retinal plane with minimally overlapping receptive fields on a grid similar to that of Fig. 2.1. However, in their model, Sandini *et al*, choose to begin this sampling at a certain distance from the

origin which is essentially the boundary of the fovea. Within this boundary, all pixels have identical minimal area. They essentially remove the singularity at the origin by dividing up the  $z$  plane into distinct regions which map to separate output images.

This same compromise is adopted by Bolduc and Levine [4] to deal with the singularity, however their model attempts a more accurate representation of the biological data. Based on the work of Wilson [14], the Bolduc and Levine model samples the input image with circularly symmetric receptive fields which have some finite amount of overlap between them. This feature accurately models data reported by Hubel and Wiesel [7] on the variation of receptive field size with distance from the fovea in Macaque Monkeys.

The foveated sensor described in this thesis is designed according to a mathematical model which borrows various features of the ones discussed above. The model attempts to implement the original  $\ln z$  mapping first proposed by Schwartz. To solve the problem of the singularity at the origin, it adopts the solution present in both the Sandini and the Bolduc and Levine models of dividing up the sensor plane into two distinct regions. Like the Wilson and the Bolduc and Levine models, it mimics the circularly symmetric receptive fields known to be present in the original biological systems. However, the fact that these receptive fields are to be formed directly in silicon as photosensitive regions, precludes any overlap between them. Therefore, the model — which will be referred to as the *Hybrid Model* — can be thought of as a combination of all of those discussed previously, with various concessions made to limits imposed by a physical implementation.

## 2.4 The Hybrid Model

The basic topology of the Hybrid model is shown in Fig. 2.2. In the sensor plane, it consists of two separate image sensing arrays. The central array is termed the *fovea* and consists of a uniform distribution of photosensitive elements whose size is fixed. As illustrated in Fig. 2.2, image points in the original input image which fall within this region are mapped directly to an output image known as the *fovea image*.

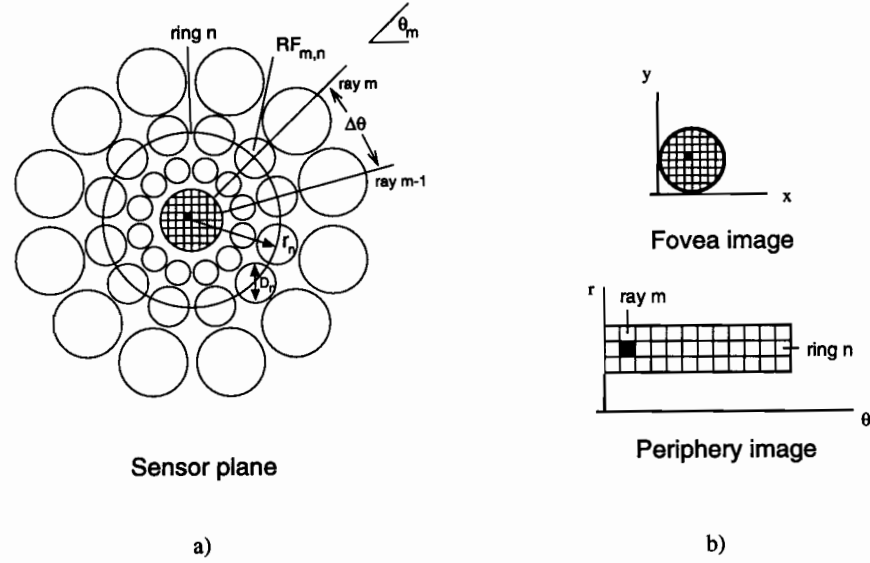


Figure 2.2: The Hybrid model. a) Illustration of the sensor plane. b) Illustration of corresponding output images. Points near the center of the input image map directly to the fovea image. Points outside the fovea map according to Eq. (2.2) to the periphery image.

The outer array is referred to as the *periphery* and is composed of photosensitive elements of varying size known as receptive fields (RF) which are distributed on a log-polar grid. The periphery consists of  $N$  rings and  $M$  rays of pixels. All pixels on ring  $n$  have the same diameter,  $D_n$ , and are located at the same radial distance,  $r_n$ , from the center of the sensor. All pixels on ray  $m$  have the same angular displacement  $\theta_m$ . Adjacent rays of pixels are separated by angular displacement,  $\Delta\theta$ . The periphery performs a mapping from cartesian to log-polar coordinates and therefore accomplishes the  $\ln z$  part of the Hybrid mapping.

#### 2.4.1 Derivation of defining equations

As the basis for a physical device, the Hybrid model must provide defining equations which can be used to locate the various photosensing elements to correctly accomplish the required image transformation. These equations are developed by working backwards from the  $w$  plane of the  $\ln z$  mapping while taking care to satisfy certain constraints inherent in a physical implementation.

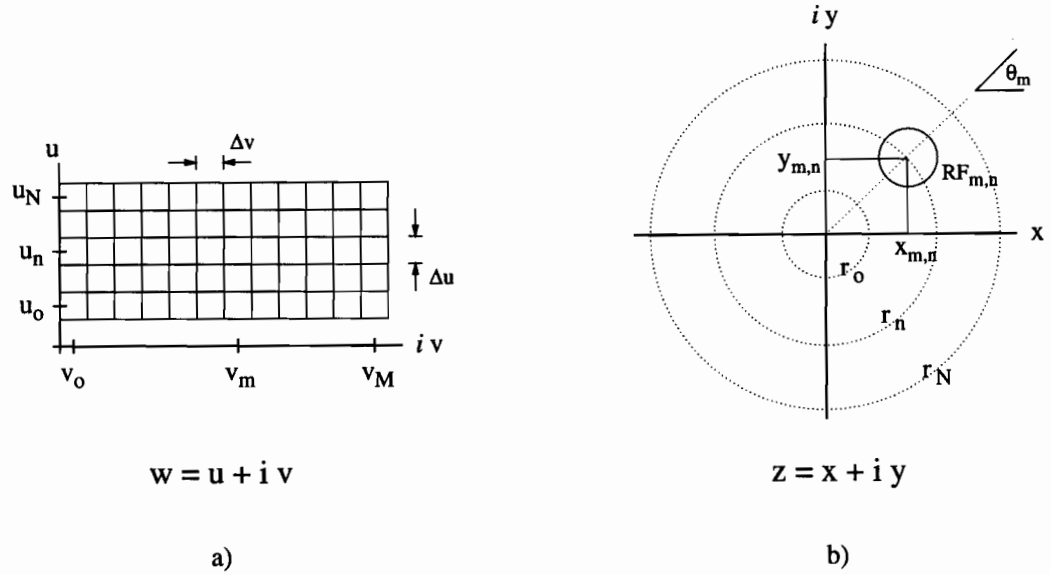


Figure 2.3: Relationship between  $w$  and  $z$  planes under the  $\ln z$  mapping. a) The  $w$  plane. b) The  $z$  plane.

Fig. 2.3a shows the  $w$  plane of the  $\ln z$  mapping for which

$$w = u + i v \quad (2.4)$$

Fig. 2.3b shows a simplified diagram of the sensor image plane which corresponds to the  $z$  part of the transformation, for which

$$z = x + i y \quad (2.5)$$

The first concession made to simplify layout of the sensor is the one which was adopted by Sandini and by Bolduc and Levine as well: the singularity at the origin is removed by instituting the following constraint,

$$u \geq u_0 \quad (2.6)$$

As explained previously, image data lost to the periphery due to this lower bound is

output to the fovea part of the mapping instead.

To satisfy the finite size of a physical image sensor, the  $w$  plane is given the following upper bounds on both axes,

$$\begin{aligned} u &\leq u_N \\ v &\leq v_M \end{aligned} \quad (2.7)$$

The use of individual photosensing devices effectively quantizes the  $w$  plane into rows and columns of pixels of width,  $\Delta v$  and height  $\Delta u$  as follows,

$$\begin{aligned} u_n &= u_0 + n\Delta u, n = 0, 1, \dots, N-1 \\ v_m &= v_0 + m\Delta v, m = 0, 1, \dots, M-1 \end{aligned} \quad (2.8)$$

Therefore, the upper boundaries of the constrained mapping become,

$$\begin{aligned} u_N &= u_0 + N\Delta u \\ v_M &= v_0 + M\Delta v \end{aligned} \quad (2.9)$$

Values for  $\Delta u$  and  $\Delta v$  are specified during the design of the periphery of the sensor as detailed below.

As discussed above, the Hybrid model makes use of circularly symmetric receptive fields in the sensor plane. As illustrated in Fig 2.2b,  $RF_{m,n}$  on ring  $n$  and ray  $m$ , is located at cartesian coordinates,  $(x_{m,n}, y_{m,n})$ . Assuming that  $RF_{m,n}$  in the  $z$  plane maps points to the corresponding pixel coordinate  $(v_m, u_n)$  in the  $w$  plane, these coordinates can be found as follows,

$$\begin{aligned} x_{m,n} + iy_{m,n} &= e^{u_n + iv_m} \\ &= e^{u_n} (\cos v_m + i \sin v_m) \end{aligned} \quad (2.10)$$

Separating real and imaginary parts, and equating  $v_m$  with  $\theta_m$ , results in the following relations,

$$\begin{aligned}x_{m,n} &= e^{u_n} \cos \theta_m \\y_{m,n} &= e^{u_n} \sin \theta_m\end{aligned}\tag{2.11}$$

As will be seen in Chapter 7, Eq. (2.11) can be used by a computer program to correctly place  $RF_{m,n}$  during physical layout of the periphery of the sensor.

To parametrize the Hybrid model, it is more convenient to use a polar coordinate system. Referring once again to Fig. 2.3, the radial displacement,  $r_n$ , is given by,

$$r_n = \sqrt{x_{m,n}^2 + y_{m,n}^2}\tag{2.12}$$

Substituting, Eq. (2.11) shows that,

$$\begin{aligned}r_n &= e^{u_n} \sqrt{\cos^2 \theta_m + \sin^2 \theta_m} \\&= e^{u_n}\end{aligned}\tag{2.13}$$

Substituting Eq. (2.8) into Eq. (2.13) gives,

$$r_n = e^{u_0} e^{n\Delta u}\tag{2.14}$$

Defining the parameters,

$$\begin{aligned}\alpha &\equiv e^{u_0} \\ \beta &\equiv 1/\Delta u\end{aligned}\tag{2.15}$$

yields the following relation for the radius of ring  $n$ ,

$$r_n = \alpha e^{n/\beta}\tag{2.16}$$

As illustrated in Fig. 2.4,  $RF_{m,n}$  is located on ray  $m$  which is at angular displacement  $\theta_m$ . The angular distance between two adjacent rays  $\theta_m$  and  $\theta_{m-1}$  in a sensor matrix containing  $M$  rays is given by,

$$\Delta\theta = \frac{2\pi}{M}\tag{2.17}$$

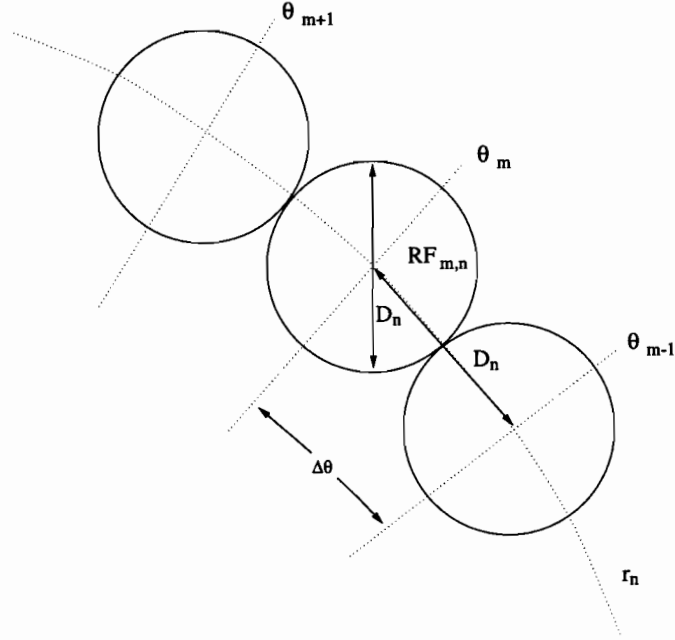


Figure 2.4: Determination of the diameter of  $RF_{m,n}$  for contiguous rays.

and is equivalent to the quantization size  $\Delta v$  in the  $w$  plane.

Assuming that all RF's on ring  $n$  are contiguous, the length of the arc  $D_n$  between rays at displacements  $\theta_m$  and  $\theta_{m-1}$  is given approximately by,

$$D_n = \Delta\theta r_n \quad (2.18)$$

This is also the diameter of all RF's on ring  $n$  for non-overlapping rays of adjacent receptive fields.

As will be seen in Chapter 7, fixing the parameters  $\alpha$ ,  $\beta$  and  $\Delta\theta$  creates a unique foveated sensor configuration within the Hybrid model, and allows for computerized layout of the periphery array using the above defining equations.

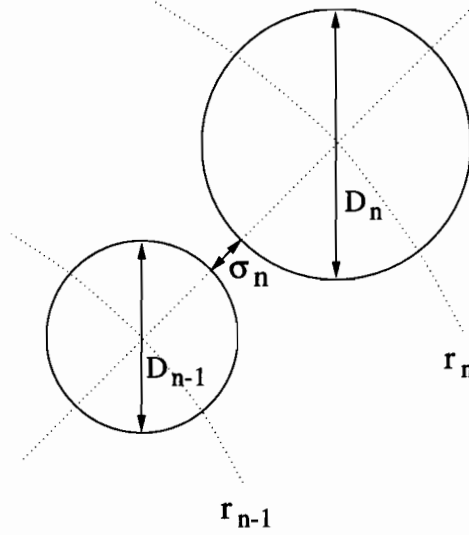


Figure 2.5: Determination of separation factor,  $\sigma_n$ , for RF's on ring  $n$  of radius  $r_n$ .

#### 2.4.2 Consequences of non-overlapping receptive fields

The model defined above is intended to be used for a sensor with non-overlapping receptive fields. It is of course conceivable that combinations of  $\alpha$ ,  $\beta$ , and  $\Delta\theta$  exist for which some overlap among adjacent rings of pixels does indeed occur.

To investigate this possibility, a separation factor  $\sigma_n$  representing the distance between two adjacent rings is defined. Referring to Fig. 2.5, the factor is given by,

$$\sigma_n = (r_n - D_n/2) - (r_{n-1} + D_{n-1}/2) \quad (2.19)$$

Substituting in Eq. (2.18) into Eq. (2.19), gives,

$$\sigma_n = r_n (1 - \Delta\theta/2) - r_{n-1} (1 + \Delta\theta/2) \quad (2.20)$$

For non-overlapping consecutive rings which are just touching the condition  $\sigma_n = 0$  is imposed, which combined with Eq. (2.20) yields the following relation,



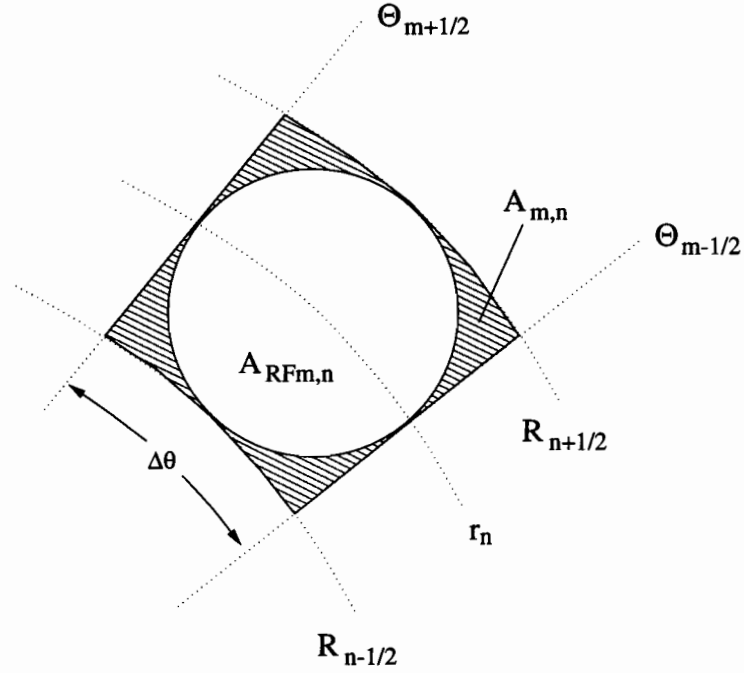


Figure 2.6: Determination of percentage of lost image data  $\gamma_i$  for  $RF_{m,n}$ .

$$\frac{r_n}{r_{n-1}} = \frac{1 + \Delta\theta/2}{1 - \Delta\theta/2} \quad (2.21)$$

which can be simplified, by substituting in Eq. (2.16) yielding,

$$e^{1/\beta} = \frac{2 + \Delta\theta}{2 - \Delta\theta} \quad (2.22)$$

or simply,

$$\beta = \frac{1}{\ln \frac{2+\Delta\theta}{2-\Delta\theta}} \quad (2.23)$$

Therefore, although  $\beta$  may be specified independently of  $\Delta\theta$ , imposing the additional constraint of Eq. (2.23) will guarantee that adjacent rings in the periphery are contiguous.

Assuming that adjacent rings are indeed contiguous, it is useful to know how much image information is lost due to the use of non-overlapping receptive fields. Referring to Fig. 2.6, the area  $A_{m,n}$ , of the polar rectangle defined by ray boundaries,  $\Theta_{m-1/2}$

and  $\Theta_{m+1/2}$ , and ring boundaries  $R_{n-1/2}$  and  $R_{n+1/2}$ , is given approximately by,

$$A_{m,n} = r_n (\Delta\theta r_n) \Delta\theta \quad (2.24)$$

and the area  $A_{RF_{m,n}}$ , of  $RF_{m,n}$  occupying the same location in the sensor matrix is given by,

$$A_{RF_{m,n}} = \pi \left( \frac{\Delta\theta r_n}{2} \right)^2 \quad (2.25)$$

Therefore, the percentage  $\gamma_i$  of image data lost may be determined as follows,

$$\begin{aligned} \gamma_i &= \frac{A_{m,n} - A_{RF_{m,n}}}{A_{m,n}} \\ &= \frac{\Delta\theta^2 r_n^2 - \pi \Delta\theta^2 r_n^2 / 4}{\Delta\theta^2 r_n^2} \end{aligned} \quad (2.26)$$

$$(2.27)$$

which reduces to,

$$\begin{aligned} \gamma_i &= \frac{4 - \pi}{4} \\ &= 0.21 \end{aligned} \quad (2.28)$$

and is independent of ring number and consequently identical over the entire array.

Therefore, the use of the Hybrid model with the constraint  $\sigma_n = 0$  results in a net loss of approximately 1/5 of the original image information, regardless of other parameters. As will be seen in Chapter 7, this is a small price to pay for the considerable simplification in layout complexity afforded by the use of circularly symmetric non-overlapping receptive fields.

## 2.5 Implementations

There have been a number of implementations of foveation algorithms documented in the literature. These vary in the models they are based on, the implementation

strategy which they adopt, as well as the performance they obtain.

Bederson, Wallace and Schwartz [15] implemented the  $\ln(z + a)$  mapping using dedicated DSP engines which averaged image data sampled from a commercially available area imager. The system outputs a logmap of 1378 pixels at a maximum rate of 30 frames/second.

Kreider *et al* [3] fabricated a fully custom CCD implementation of the model described by Sandini. The device contains a periphery array composed of 30 rings and 64 rays of photoelements and a fovea array containing 102 pixels.

Finally, Bolduc and Levine [4] implemented their overlapping receptive field model on a network of TMS320C40 DSP processors. This system produces separate fovea and periphery images of variable size according to biologically-motivated parameters set by the user. The system operates at a minimum rate of 10 frames/second on an input image of 484 x 484 pixels.

For the purpose of illustration, examples of input and output images from this system are shown in Fig. 2.7. The input image consists of a synthetic pattern designed to highlight the features of the foveated mapping. The periphery output image shows how straight lines are mapped to connected curves by the algorithm. The fovea output image shows that features constrained to the central part of the image are reproduced faithfully at the original resolution of the input image. The starting resolution is 484 x 484 pixels, while the periphery output resolution is 20 x 50, and the fovea output is 101 x 101. This constitutes a reduction of data of approximately 20:1. These numbers illustrate the considerable savings which may be realized through the use of foveation.

## 2.6 The use of foveation in robotic systems

At present the most enticing feature of the foveated mapping is the tremendous gain in data compression which it affords. Schwartz estimates that a visual cortex capable of processing a uniform one-to-one mapping of every point in the retinal image would have to be three to four orders of magnitude larger (surface area) than the one which

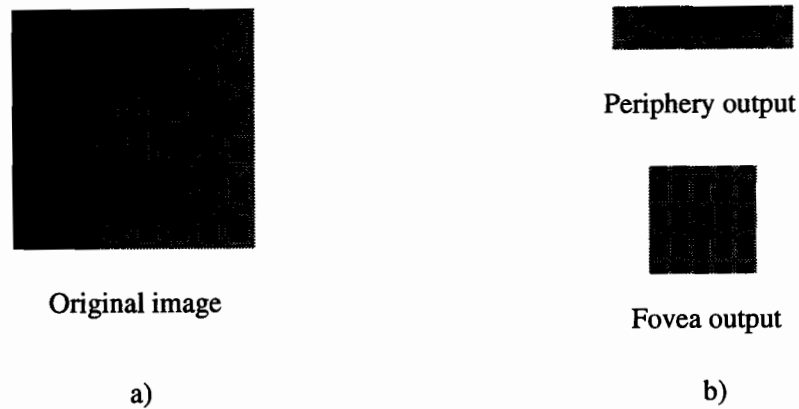


Figure 2.7: Example images from the Bolduc and Levine model. a) Original image. 484 x 484 pixels. b) Output images. Periphery image: 12 x 50. Fovea image: 101 x 101. (courtesy of Bolduc and Levine)

presently exists [13]. Nevertheless, humans are quite capable of functioning under the reduced information set which foveation provides. This is because we are able to move our eyes to place the fovea at the center of attention in an image.

Sandini *et al* note that under their mapping the same amount of pixel information is contained in an unmapped image as in a transformed image of an original scene 30 times larger than the non-mapped image. This is a significant result notes Sandini[11], since very often,

“execution time and computer memory space must be spent to eliminate primarily the redundant information in acquired images in order to extract the features required by the execution of an assigned task”

when this function is performed automatically by the foveated mapping. The only draw back is that the robot ‘eye’ now requires a sophisticated control and attentional algorithm to correctly position the fovea over the relevant task-related image data.

Rojet and Schwartz [9] estimate that for equivalently sized retinal sensors for which a uniform array requires 206,000 pixels, their nonlinear sensor would require merely 2,100 pixels. Further, if they double the foveal resolution, the uniform sensor now requires 823,000 pixels — an increase of 300%, whereas the log sensor now requires 2,550 pixels — an increase of only 21%. This indicates that because the complexity

of the foveated sensor is related to the log function, its pixel count grows acceptably with increase in foveal resolution.

The foveated sensor is therefore ideally suited to applications such as mobile robot vision in which the robot must be capable of responding quickly to stimulus in its environment but does not necessarily have substantial on board computational power. The resolution available to such a robot would be sufficient for execution of simple tasks.

If the robot were required to perform tasks involving image recognition it would profit from the rotation and scale invariance inherent in foveation. It is still unclear whether the biological system actually makes use of these properties to facilitate character or template recognition, however researchers have used them to advantage for this purpose.

For a robot to navigate through its environs successfully it must be able to perform estimates of the optical flow within an image. Any motion which is radial in the retinal domain is mapped to linear translation in the cortical domain. This significantly simplifies the detection of movement of the world image as the robot navigates through space, since objects naturally migrate from the center to the periphery. This feature can be used to estimate collision times as well as for depth-from-motion type calculations [9].

Therefore, it is clear that significant motivation exists for the use of foveation in robotic systems. It remains to be seen, however, whether inexpensive and useful foveated sensors can be created to facilitate the widespread use of the algorithm.

## 2.7 Summary

This chapter presented a brief introduction to the concept of the foveated mapping. Readers interested in a more detailed discussion are referred to the work of Bolduc [16] in particular, and to that of Schwartz and of Sandini. The intention here was to provide adequate background information before explaining the design of the sensor

itself. The next chapter continues along these lines by discussing the nature of light sensing in semiconductor devices.

## Chapter 3

# Silicon phototransducers

### 3.1 Introduction

One of the most important components in any mobile robot system is its link to the outside world. Whether computations are done locally using dedicated electronic hardware or remotely using a complex network of information processors, the environment must first be sensed and converted to a machine-intelligible form. In the case of computer vision this is accomplished through the use of image sensing equipment which transduces incident light to electrical signals.

This chapter provides an introduction to the use of semiconductor devices for accomplishing this task. The simplest of these devices is the *photodiode* which is examined in detail in Section 3.2. More complex phototransducers do exist, and these are examined in Section 3.3. The remainder of the chapter is devoted to a discussion of Photon Flux Integration, a method of improving the photodiode's sensitivity which makes it useful in practical applications.

### 3.2 The photodiode

The photodiode is the principal unit of solid-state imagers. This section examines photodiode operation at the physical level, and uses it to introduce some important concepts in optoelectronics.

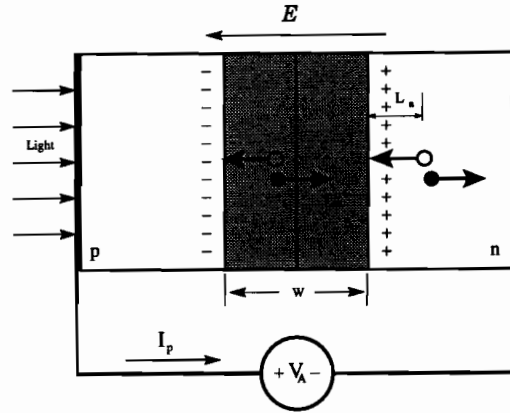


Figure 3.1: Photogeneration in a *pn* photodiode. Generated electrons and holes drift under the influence of the electric field,  $E$ , resulting in photocurrent  $I_p$ .

### 3.2.1 Photogeneration in a PN junction photodiode

Semiconductor devices are useful in image sensing applications because they are capable of transducing incident illumination through a process known as *photogeneration*. When a photon is absorbed within a region of semiconductor such as silicon, it liberates an electron from the semiconductor matrix creating a photogenerated electron-hole pair [17]. If this event occurs within a region of space charge such as that which exists at the boundary of a reverse-biased *pn* junction, the pair will be immediately separated by the potential difference which exists across the device. Therefore, a *pn* junction diode can be used as a light transducer by exposing it to an incident photon flux, and measuring the resultant flow of charge, known as *photocurrent*.

Figure 3.1 shows a two-dimensional *pn* junction diode which is exposed to incident light. The junction is reverse-biased by an external voltage source  $V_A < 0$ . The source is assumed to be connected at either end of the diode using completely transparent ohmic contacts. Near the boundary between the *p*-type and *n*-type material, there exists a depletion region of width [18],

$$W = \sqrt{\frac{2K_s\epsilon_0(V_{bi} - V_A)}{q} \frac{(N_A + N_D)}{N_A N_D}} \quad (3.1)$$



where  $K_s$  is the relative dielectric constant of the semiconductor ( $K_s$  of silicon is 11.8),  $\epsilon_0$  is the permittivity of free space ( $\epsilon_0 = 8.854 \times 10^{-14}$  Farad/cm),  $q$  is the charge on an electron ( $q = 1.602 \times 10^{-19}$  Coul.),  $V_{bi}$  is the built in junction voltage (usually about 0.7 V),  $V_A$  is the reverse-bias voltage maintained by the external voltage source,  $N_A$  is the acceptor impurity concentration (# of impurity atoms /  $\text{cm}^3$ ), and  $N_D$  is the donor impurity concentration (# of impurity atoms /  $\text{cm}^3$ ).

Photons which are absorbed within the silicon outside of the depletion region of width  $W$ , generate electron-hole pairs which may experience one of two fates. A certain number of these are annihilated through recombination and therefore do not contribute to the photocurrent flowing within the device. The remainder, however manage to diffuse into the depletion region from the bulk semiconductor and thus contribute to the photocurrent. As illustrated in Fig. 3.1, the maximum distance that these carriers are capable of traveling within the bulk semiconductor is termed the *minority carrier diffusion length*,  $L_n$  [19] and is an intrinsic quality of the semiconductor.

Those photons which are absorbed *within* the depletion region of width  $W$  generate electron-hole pairs which directly contribute to the photocurrent. Generated electrons drift across the junction from left to right into the  $n$ -type region. Generated holes drift from right to left into the  $p$ -type region. As indicated in Fig. 3.1, the combined drift of positive and negative charge results in a flow of photocurrent,  $I_p$ , in the external circuit.

### 3.2.2 Quantum efficiency

Not all photons incident on a semiconductor will be absorbed by the material. The sensitivity of a particular semiconductor to optical radiation is its *quantum efficiency*,  $\eta$ , and is defined as the fractional number of electron-hole pairs produced in the semiconductor for a given amount of incident photons [20],

$$\eta = \frac{\# \text{ generated electron-hole pairs}}{\# \text{ incident photons}} \quad (3.2)$$

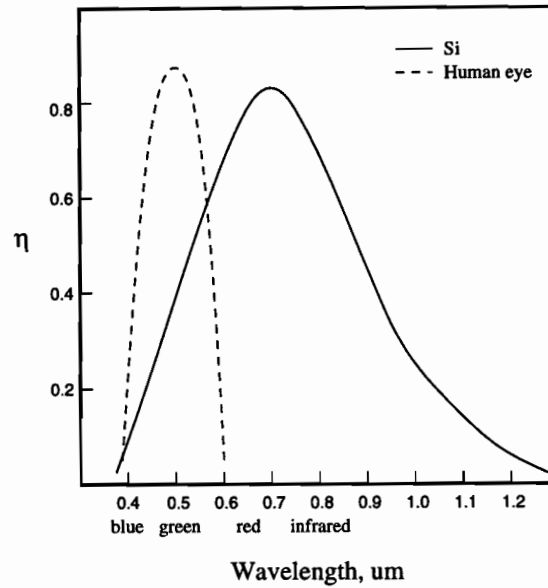


Figure 3.2: Graph of the quantum efficiency of silicon  $p^+n$  photodiodes with respect to wavelength. Also shown for comparison is a normalized plot of the response of the human eye.

Quantum efficiency is an empirical quantity specific to each photosensitive element and is a function of the wavelength of the incident radiation. A graph of  $\eta$  vs. wavelength for typical silicon photodiodes is shown in Fig. 3.2 [21]. Also shown for comparison is a normalized plot of the response of human photopic vision. It is important to note that silicon is most sensitive to wavelengths near the infrared, and rather insensitive to those which we perceive as the cooler shades of the spectrum such as blue and violet. This makes the implementation of color sensors more difficult, however a good deal of work has already been done in the field.<sup>1</sup>

### 3.2.3 Responsivity

In addition to characterization of the sensitivity of the semiconductor, it is useful to characterize the sensitivity of the phototransducer as a whole. The sensitivity of a given device is commonly referred to as its *responsivity*,  $\mathcal{R}$ .

<sup>1</sup>See for example reference [22].

The strictest definition of  $\mathfrak{R}$  is the amount of photocurrent produced by the device per unit incident optical power. Therefore, in the case of the  $pn$  junction photodiode of Fig. 3.1, the photocurrent  $I_p$  is given by,

$$I_p = \mathfrak{R}AP \quad (3.3)$$

where  $I_p$  is in Amps,  $\mathfrak{R}$  is in A/W,  $A$  is the cross-sectional area of the  $pn$  junction in  $\text{cm}^2$ , and  $P$  is the incident light irradiance in  $\text{W}/\text{cm}^2$ . In this case,  $\mathfrak{R}$  is related to  $\eta$  in the following manner,

$$\mathfrak{R} = q\eta\left(\frac{\lambda}{hc}\right) \quad (3.4)$$

where  $q$  is the charge on an electron ( $q = 1.602 \times 10^{-19}$  Coul.),  $\lambda$  is the wavelength of the incident radiation, and  $c$  is the speed of light ( $c = 3 \times 10^8$  m/sec). From Eq. (3.4), it is important to note that a given value of  $\mathfrak{R}$  is meaningless unless the wavelength of the incident radiation is specified. The quantity in Eq. (3.4) in brackets is the energy of an incident photon of wavelength  $\lambda$ . Therefore, Eq. (3.4) shows that incident photons bearing a certain amount of energy are absorbed in the semiconductor according to  $\eta$  and converted to a photocurrent.

#### 3.2.4 Dark current

In addition to optical means, generation in a reverse-biased  $pn$  junction takes place by thermal excitation of the crystal lattice [18]. Thermally generated electron-hole pairs contribute a significant portion of the reverse current in a photodiode which is not subject to illumination. Therefore the minimum current flowing in a reverse-biased photodiode is strongly temperature dependent and is termed the *dark current*,  $I_s$ , of the device.

As will be noted in Chapter 6, the dark current,  $I_s$ , is one of the principal sources of spatial and temporal noise in solid-state imagers. It increases by a factor of 2 for every 10-degree Celsius increase in temperature, and is a Poisson-distributed random process whose mean varies across the substrate [23]. Therefore, in applications requiring extremely sensitive optical measurements, such as astronomy, image sensors

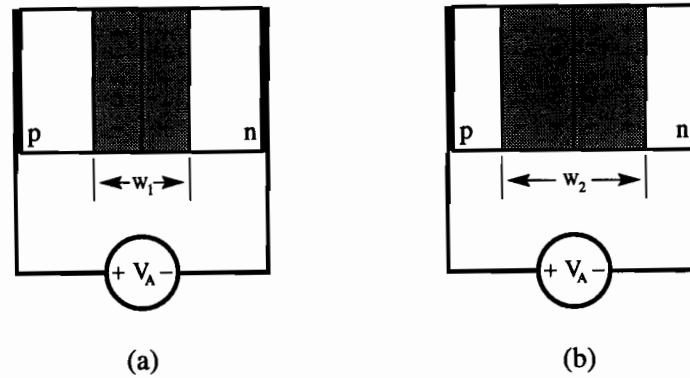


Figure 3.3: Reverse-biased  $pn$  diode junction capacitance. a) The presence of reverse-bias voltage  $V_A < 0$  generates a space charge region of width  $W_1$ . b) Increasing the reverse-bias, causes the region to grow to width  $W_2$  by adding charge to either side of the junction capacitor.

are often cooled to reduce the effects of thermal noise on the image.

### 3.2.5 Reverse-biased diode capacitance

As will be seen later, a very important property of the depletion region within a reverse-biased  $pn$  photodiode is its ability to retain charge and thus behave as a capacitor. Fig. 3.3 illustrates this point. According to Eq. (3.1), the width of the depletion region is proportional to the square root of the voltage,  $V_A$ , across the  $pn$  junction.

If  $V_A$  is made more negative, the depletion region will grow by removing majority carriers from both the  $n$  and  $p$  sides of the device. On the  $n$  side, electrons are stripped away from their respective donor atoms, uncovering positive charge. On the  $p$  side, holes leave behind negatively charged acceptor ions. Therefore, the process of widening the depletion region by making  $V_A$  more negative implies the *addition* of charge on either side of the  $pn$  junction. Similarly, the process of shrinking the depletion region by making  $V_A$  less negative implies the *removal* of charge from either side of the  $pn$  junction.

This type of behaviour is reminiscent of a parallel plate capacitor. In direct analogy, the capacitance of the reverse-biased  $pn$  junction can be defined as,

$$C_J = \frac{K_s \epsilon_0 A}{W} \quad (3.5)$$

where  $A$  is the cross sectional area of the junction,  $W$  the width, and  $K_s \epsilon_0$  the dielectric constant of the semiconductor. Substituting Eq. (3.1) into Eq. (3.5),

$$C_J = \frac{K_s \epsilon_0 A}{\sqrt{\frac{2K_s \epsilon_0 (V_{bi} - V_A)}{q} \frac{(N_A + N_D)}{N_A N_D}}} \quad (3.6)$$

shows that the capacitance can be directly determined once the reverse-bias voltage is known.

In the case where  $V_A$  is composed of a small signal component plus a constant bias voltage, the device can be assumed to behave more or less as a linear capacitance with value given by Eq. (3.6). However, for larger variations in  $V_A$ , the response becomes decidedly nonlinear. Although, this leads to distortion in applications requiring linear capacitance such as signal processing, the problem is less severe for solid-state imaging. As will be shown later, the nonlinear reverse-biased diode capacitance plays an important role in modern sensors.

### 3.2.6 Photodiode schematic model

Owing to the above described characteristics, the photodiode is typically represented as a combination of basic schematic building blocks as illustrated in Fig. 3.4a. The current source is used to model the contributions of the photocurrent  $I_p$ , and the dark current  $I_s$ , to the reverse-bias current. In most cases,  $I_s$  is assumed to be negligible as compared to  $I_p$  and therefore it is omitted. The capacitance  $C_p$  is used to model the reverse-biased junction capacitance  $C_J$  of the photodiode. In some cases, the symbol in Fig. 3.4b will be used for simplicity and is understood to represent the model of Fig. 3.4a.

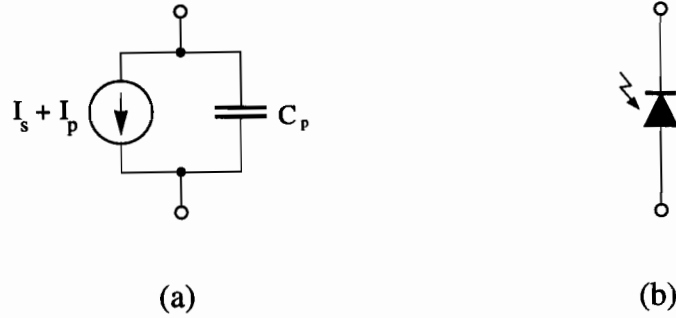


Figure 3.4: Photodiode schematic model. For simplicity the network in a) is understood to be represented by the symbol in b).

### 3.3 Other photosensitive devices

Although the photodiode is ubiquitous in image sensing circuitry, other devices do exist and may be found in specialized applications. It is worth examining a few of these here, since they will be mentioned later on in the discussion.

#### 3.3.1 The phototransistor

Even before photodiodes became popular in MOSFET-scanned arrays, the *phototransistor* was being considered as a candidate for image sensing applications [24]. As illustrated in Fig. 3.5a, the phototransistor is simply a Bipolar Junction Transistor (BJT) whose base is left floating. The current flowing in the emitter,  $I_E$ , is  $(\beta + 1)$  times that in the base,  $I_B$ . If  $I_B$  is a photocurrent, the output of the device,  $I_E$ , will be an amplified version of the incident light signal. Since the value of  $\beta$  can vary between 50 and 300 depending on the width of the base region [25], significant gain may be realised.

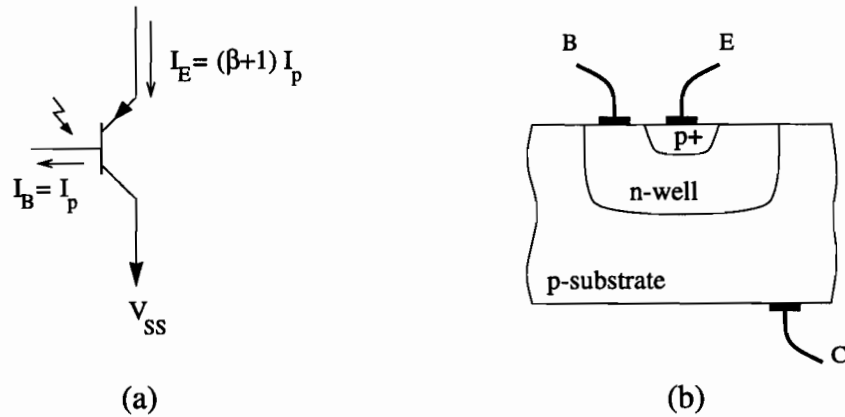


Figure 3.5: The CMOS *pnp* phototransistor. a) Photocurrent flowing out of the base is amplified at the emitter. b) The transistor is formed by implanting a  $p^+$  emitter diffusion in an  $n$ -well on a  $p$ -type silicon substrate.

Although BJTs used in high quality analog microelectronics require special processing steps, adequate phototransistors can be implemented in a standard CMOS process. These transistors exist as a natural, often detrimental<sup>2</sup> byproduct of the fabrication process, and are therefore termed *parasitic transistors*.

Figure 3.5b shows a cross section of a CMOS *pnp* phototransistor. The transistor is created by forming a  $p^+$  diffusion inside an  $n$ -well on a  $p$ -type substrate. The  $p^+$  diffusion is the emitter of the device, and the  $n$ -well is its base. The  $p$ -type substrate forms the collector of the transistor. Photons which are incident on the  $pn$  junction at the edges of the  $n$ -well generate carriers which are amplified by the transistor and result in a current at the  $p^+$  emitter contact.

The main objection to the use of phototransistors in CMOS image sensors is their size. The requirement for isolation in the CMOS process places restrictions on the

<sup>2</sup> "latchup" is a potentially disastrous situation which can occur in CMOS microelectronics. Cross-coupled, parasitic BJT's "latch" into the on state and conduct very large currents to the substrate. Aside from rendering the effected gates inoperative, latchup can permanently damage the semiconductor due to the excessive heat generated. The threat of latchup can be avoided through careful design practise.

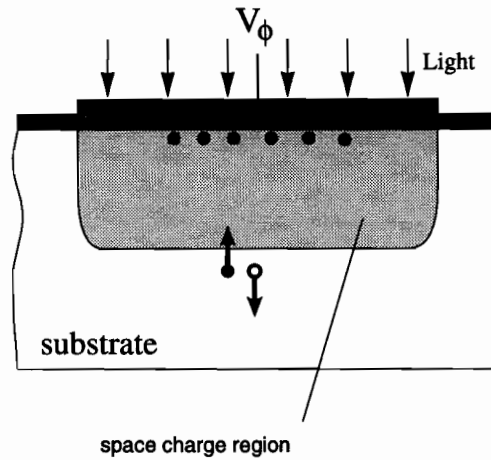


Figure 3.6: The MOS-C photogate. A space charge region is created below the gate by bringing  $V_\phi$  above a threshold. Photogenerated electrons are swept up and collect just below the gate forming a charge packet proportional to the incident light irradiance.

minimum width and separation of device wells. For example in Nortel's CMOS4S process,  $n$ -wells must be separated by a minimum of  $13\ \mu\text{m}$ , and must be at least  $3\ \mu\text{m}$  wide [26]. Therefore, phototransistors are too expensive for high resolution applications. However, they remain useful in special cases.<sup>3</sup>

### 3.3.2 The MOS-C

As was seen in Section 3.2.1, one of the most important phenomena for transduction of incident light in photodiodes is the existence of a space charge region to separate photogenerated electron-hole pairs. Although the existence of a physical  $pn$  junction is useful for this purpose it is not at all necessary.

Charge Coupled Device (CCD) image sensors, which will be examined in the next chapter, make use of a photosensitive device with a voltage induced space charge region. This device, known as a Metal Oxide Semiconductor-Capacitor (MOS-C)

<sup>3</sup>See for example reference [27].



*photogate*, is illustrated in Fig. 3.6. A transparent gate electrode is formed on the surface of a  $p$ -type (or  $n$ -type) substrate. When the voltage  $V_\phi$  is increased beyond a certain threshold, the semiconductor immediately below the gate is completely depleted of majority carriers: a situation known as *deep depletion* or *inversion*. This creates a thin region of space charge below the gate [28].

Light incident from above passes through the transparent electrode into the semiconductor and generates electron-hole pairs. As was the case for the  $pn$  photodiode, charge carriers within a distance  $L_n$  can diffuse into the depletion region. When this happens, the electron hole pairs are separated. Electrons drift up and collect right below the gate; holes drift into the bulk semiconductor. The net effect is the formation of a charge packet below the gate in direct proportion to the incident photon flux on the device.

### 3.3.3 The subthreshold MOSFET

When a MOSFET is forced to conduct very low amounts of current ( $\leq 100nA$ ), it enters into the *subthreshold* region of operation [25]. In this region, the normal square law relationship between drain current and gate voltage becomes exponential,

$$I_D = I_0 e^{-(V_S - V_{S0})/V_T} \quad (3.7)$$

where  $I_D$  is the drain current,  $I_0$  is the drain current at the onset of threshold,  $V_S$  is the source voltage,  $V_{S0}$  is the source voltage at the onset of threshold, and  $V_T$  is the thermal voltage (at room temperature  $V_T = 0.026$  V) [29]. Rearranging Eq. (3.7),

$$V_S = V_{S0} - V_T \ln \frac{I_D}{I_0} \quad (3.8)$$

shows that the variation in voltage at the source of the MOSFET is proportional to the natural logarithm of the drain current.

Therefore, by forcing a MOSFET to carry the low current of a photodiode, a special photosensitive device is created whose output voltage varies with the log of the photocurrent. As will be seen in the next chapter, this device compresses the dynamic range of the incident light signal, and is therefore operable over a wide range

of lighting conditions.

### 3.3.4 The CMD

Researchers in Japan [30] and the US [31] have recently introduced a new photo-sensor called a Charge Modulation Device (CMD). This is essentially a MOSFET operated in a similar fashion to a MOS-C photogate (Section 3.3.2). Photocharge is collected below the transparent MOSFET gate. This charge *modulates* the channel below the gate of the MOSFET and therefore alters the current flow in the device. The advantage of such an arrangement is light-dependent current amplification with nondestructive readout.

Several imaging systems based on this concept have already been developed [32], [31], and the results are promising. However, due to the complex method of operation of this new technology it does not appear suitable for general purpose CMOS vision sensors, and will likely remain the exclusive domain of very high resolution sensor applications.

## 3.4 Photon flux integration

The majority of successful image sensing devices to date have made use of photodiodes operating in the *integration mode* (sometimes referred to as *Photon Flux Integration*), outlined by Weckler [33]. This section reviews this mode of operation and shows why it is the most effective way to use photodiodes for image sensing applications.

### 3.4.1 Integration mode

As stated in Section 3.2.3, a photodiode receiving incident radiation  $P$ , will generate a photocurrent  $I_p$ . If  $I_p$  is assumed to be relatively constant in time (which is true for image sensing and will be shown below), then over a given period  $t_i$ , the photodiode will generate a total charge,

$$Q_p = I_p \times t_i \quad (3.9)$$

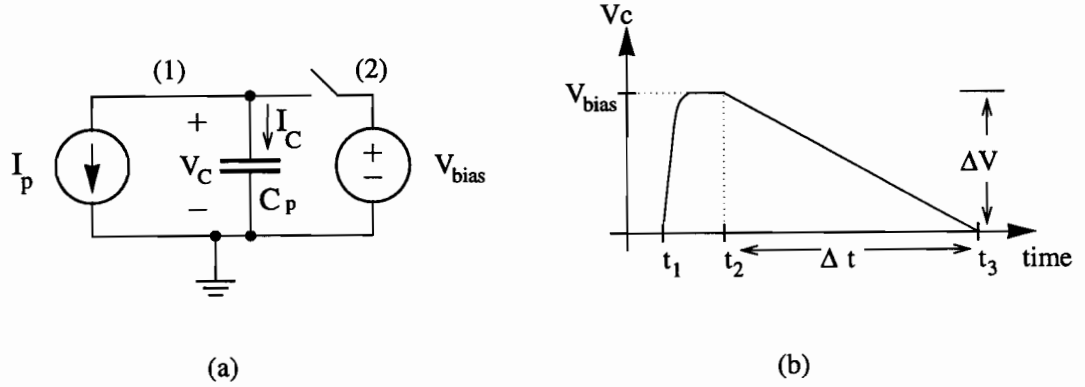


Figure 3.7: Photon flux integration. a) The photodiode is charged to  $V_{bias}$  by shorting nodes (1) and (2). b) The diode voltage  $V_c$  decays due to the constant photocurrent,  $I_p$ .  $\Delta V_C$  is the light dependent signal voltage, and  $\Delta t$  is the period of integration.

in response to the optical radiation. If this charge is stored in a capacitor  $C$  it will be automatically converted to a voltage,

$$V_p = \frac{Q_p}{C} \quad (3.10)$$

proportional to the incident photon flux. Through integration over time, the initially weak, noisy photocurrent  $I_p$  will be amplified and converted into a more manageable quantity.

Figure 3.7 shows one possible way of accomplishing photon flux integration. The current source  $I_p$  represents current being drawn as a result of photons impinging on the  $pn$  junction, and the capacitor,  $C_p$ , represents the diode parasitic capacitance (Section 3.2.5). The voltage source,  $V_{bias}$ , is used to set a known starting point for the integration.

At time  $t = 0$  the switch is open. The capacitor is completely depleted of charge and therefore the voltage across it,  $V_c$ , is 0 V. At time  $t_1$  the switch is closed connecting the voltage source to the rest of the circuit. Charge is transferred from the

voltage source to the capacitor until the condition  $V_c = V_{bias}$  is satisfied. At a certain time  $t_2$ , with the capacitor fully charged, the switch is opened. The capacitor and the current source are isolated. Due to the incident radiation a constant current  $I_p$  flows from node (1) to ground. However by Kirchhoff's current law (KCL), the total current entering node (1) must be equal to the total current leaving. Therefore, the capacitor current  $I_c$  is given by,

$$I_c = -I_p \quad (3.11)$$

From circuit analysis theory,  $I_c$  and  $V_c$ , are related by the following equation,

$$I_c = C_p \frac{dV_c}{dt} \quad (3.12)$$

rearranging Eq. (3.12), and substituting Eq. (3.11),

$$dV_c = -\frac{1}{C_p} I_p dt \quad (3.13)$$

Integrating both sides from the start of photon integration,  $t_2$ , until the end,  $t_3$ , gives,

$$\begin{aligned} \int_{t_2}^{t_3} dV_c &= \frac{-1}{C_p} \int_{t_2}^{t_3} I_p dt \\ V_c(t_3) - V_c(t_2) &= -I_p \frac{(t_3 - t_2)}{C_p} \\ \Delta V_c &= -\frac{\Delta t}{C_p} I_p \end{aligned} \quad (3.14)$$

Equation (3.14) relates the total change in voltage on the capacitor,  $\Delta V_c$ , to the photocurrent  $I_p$ . Therefore,  $\Delta V_c$  is taken to be an amplified version of  $I_p$ , with gain equal to  $\frac{\Delta t}{C_p}$ . The interval,  $\Delta t$ , is termed the *period of integration*, and can be altered to directly adjust the gain of the system. Longer integration times will allow the detection of minute light levels, whereas short integration times make accurate measurement of bright light feasible.

### 3.4.2 Exposure

Photodetectors operated in the integration mode require a new definition of  $\mathfrak{R}$  to account for the temporal nature of their sensitivity. A measure of the number of photons collected through photon flux integration is the *exposure*,  $E$ , given by,

$$E = P \times \Delta t \quad (3.15)$$

where  $P$  is the incident light irradiance in  $\text{W}/\text{cm}^2$ . Examining Eq. (3.15), it is apparent that  $E$  has units of  $\text{W}/\text{cm}^2 \times \text{time}$ , or  $\text{J}/\text{cm}^2$ . Therefore, the definition of  $\mathfrak{R}$  for integrating devices is,

$$\mathfrak{R} = \frac{\text{Output Voltage}}{E} \quad (3.16)$$

with units of  $\text{V}/(\text{J}/\text{cm}^2)$  (or in practise,  $\text{V}/(\mu\text{J}/\text{cm}^2)$ ). This is the most common sensitivity measure employed to characterize CCD image sensors, and will be used here to characterize CMOS imagers as well.

## 3.5 Summary

In this chapter, some of the most important concepts in the use of semiconductor phototransducers were examined. The photodiode is the most popular of these devices in use today for image sensing applications. It was examined in detail in Section 3.2 where a convenient first order model of its operation was explained. As will be seen in Chapter 5, this model is very useful for simulation of CMOS imager arrays. Although, the photodiode is the most popular transducer, other devices do find use in specialized applications. In particular the MOS-C introduced in Section 3.3 will be seen again in the next chapter in the context of a discussion of CCD imaging arrays. Finally, as will be seen in the following chapters, the use of photon flux integration as described in Section 3.4 is essential to the use of photodiodes in practical imaging applications. In the next chapter some of these applications will be examined as the basic phototransducer is extrapolated to complete image sensing arrays.

## Chapter 4

# Solid-state image sensing

### 4.1 Introduction

With the publication of the theory behind the photon flux integration mode in 1967, solid-state image sensing research began to accelerate at a rapid pace. By the end of the decade, various private and government-funded labs had produced functioning two dimensional image sensing arrays based on this approach [34], [24], [35]. Whether motivated by aerospace or commercial applications, the principal goal was the same: research centered around the realization of a transistor-scanned photodiode camera capable of meeting or exceeding the required NTSC television standard of 525 vertical lines [36].

This chapter first defines the image sensing problem in Section 4.2 and then provides a discussion of early attempts at solid-state imaging in Section 4.3. These efforts were soon eclipsed by the Charge Coupled Device (CCD) which within three years of its initial development at Bell Laboratories, was already well on its way to becoming the technology of choice for image sensing [37]. The elementary operation of the CCD is examined in Section 4.4 and shown to be the source of both its success and its inevitable decline.

Finally, the state of the art in image sensing technology for computer vision applications is presented. In the past few years there has been an explosion in the proliferation of custom integrated circuits uniting the previously separate tasks of image sensing and image processing. These circuits all share the promise of immense

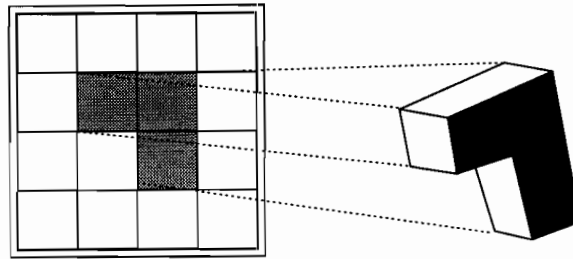


Figure 4.1: An object is projected onto the sensor plane.

savings in system mass, cost and speed of operation. Sections 4.5 and 4.6 examine the promise of CMOS image sensing technology for future computer vision applications.

## 4.2 The image sensing problem

To motivate the discussion of solid-state imagers, it is useful to examine the nature of the image sensing problem. The most important issue to be dealt with in any image sensor is that of information access. As illustrated in Fig. 4.1, light bearing information about the object to be imaged, is projected onto the sensor plane. The sensor is divided up into an array of discrete elements termed *pixels*. Each pixel transduces and stores the light information incident at that location in the matrix. The problem then becomes how to efficiently access the large amount of parallel information contained within the array. As will be seen in the Section 4.4, one way is to use a type of ‘bucket-brigade’ approach in which pixels are shifted out in quick succession, using an analog shift register. Alternately each pixel could be accessed independently, in a manner similar to Random Access Memory (RAM) as is the case with the image sensor described in Chapter 5.

### 4.3 First generation solid-state imagers

The first useful solid-state imaging devices to come out of research labs in the late 1960's consisted of arrays of photodiodes integrated with selection circuitry. When these so-called MOSFET-scanned imagers were first introduced, integrated circuit technology was still in its infancy. Therefore, the resolution of initial arrays was quite inadequate to meet the goal of a solid-state camera. For this reason, MOSFET arrays were soon abandoned in favor of the more promising CCD imager technology. Nevertheless, some low resolution MOSFET imagers were manufactured for special purpose applications.

One of the only arrays ever to go into large scale production was developed by Koike *et al* in 1980 [22]. This was a color imager with  $484 \times 384$ . Other implementations of MOSFET-scanned sensor technology included a low power, portable reading aid for the blind [38], 35mm camera electronic autofocus systems [39], and digital card readers [35]. At present, at least one company still produces simple, MOSFET-scanned photodiode arrays for general purpose applications [40].

As will be seen in Section 4.6.3, MOSFET-scanned arrays form the basis of the most common modern-day CMOS imaging technology. This same technology is integral to the design of the CMOS foveated sensor and will therefore be examined in detail in the next chapter.

### 4.4 Charge coupled imagers

Although solid-state image sensing began with MOSFET-scanned diode array implementations, the push to break the NTSC resolution barrier soon led researchers to abandon the strictly MOSFET based approach. With the announcement in 1970 [41] of the development of a totally new semiconductor device, the era of Charge Coupled Device (CCD) imagers had begun.



#### 4.4.1 Charge coupling

The basic concept behind CCD imagers is that of charge coupling. The technique, as first proposed by Boyle and Smith [41], is relatively simple and compact, and can lead to the very densely packed architectures necessary for high resolution image sensing.

In order to explain the concept, a two pixel CCD line imager is examined. Fig. 4.2a is a cross section showing the two light sensing/charge transfer stages as well as the output stage. The CCD is composed of a series of closely spaced MOS capacitors. As explained in Chapter 3 above, the MOS-C converts incident light to a charge packet stored just under the gate of the device. Although not a photodiode, the MOS-C collects charge in much the same way as described in Chapter 3, through photon flux integration. This charge is represented in Fig. 4.2a by the dark circles underneath the  $\phi_2$  gates of stages 1 and 2. The packet forms when the  $\phi_2$  voltage is brought high, effectively creating a *potential well* in which generated electrons are collected.

In the literature an analogy is drawn between the potential well containing charge and a bucket containing water. In the case of a CCD built on a *p*-type substrate (as in Fig. 4.2), the more positive the voltage placed on a gate, the *deeper* the potential well which forms beneath it. Figure 4.2b-h illustrates the potential well concept. The amount of electrons contained in each well is illustrated by the height of the “water-line” (shaded areas in the figure). Charge transfer is accomplished by changing the heights of various potential wells in succession, effectively *pouring* electrons from one stage of the CCD to the next.

In order to sample and read out one line of optical information in this 1-D image, the gates of the CCD are clocked as illustrated in Fig. 4.2b-h. Photocharge is collected during the integration period (Fig. 4.2b), under gates  $\phi_2$  of each stage. Once integration is complete, charge is read out by shifting from left to right in rapid succession.

One complete shift right involves three cycles. As shown in Fig. 4.2c-e the first cycle moves charge from underneath gates  $\phi_2$  to gates  $\phi_3$  within a stage. The next cycle (not illustrated) moves charge from gates  $\phi_3$  in one stage to gates  $\phi_1$  in the next stage. Finally, the last cycle moves charge from gates  $\phi_1$  to  $\phi_2$  within each stage and

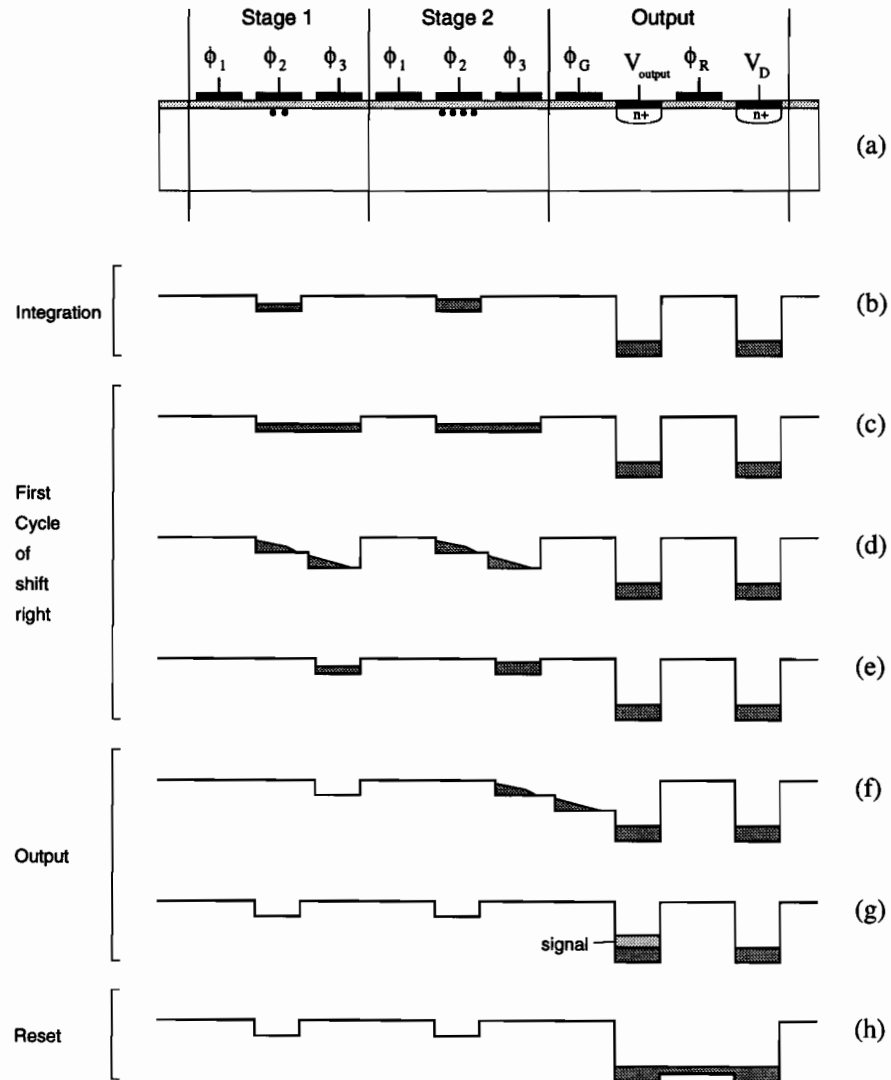


Figure 4.2: Operation of a 2-stage CCD line imager. a) Cross-section of the device. b) Integration of photocharge. c)-e) First of three cycles in a shift right one stage. f)-g) Readout of signal charge. h) Charge on the output diffusion is reset to the reference level by clocking  $\phi_R$ .

therefore ends one complete shift right of the initial optical information.

After each complete shift right of charge by a stage, information at the end of the shift register is read out by the output stage. As illustrated in Fig. 4.2a, the output stage is composed of a floating  $n+$  diffusion which is electrically connected to an output amplifier (not shown). The amount of charge present in the  $pn$  junction formed by this diffusion directly changes its capacitance and consequently the voltage  $V_{output}$ . Therefore, by first resetting this charge to a known value (Fig. 4.2h), and then shifting in a signal charge packet on top of that (Fig. 4.2g), analog charge information is converted to a voltage for subsequent processing and readout off chip.

#### 4.4.2 Charge transfer efficiency

The analog nature of the CCD delay line places strict requirements on the preservation of charge in each transfer operation. The *charge transfer efficiency*, CTE, is defined as the percentage of charge successfully transferred from one stage to the next. For a 3-phase,  $n$  stage CCD delay line, the value of an initial charge packet,  $P_0$ , after  $n$  stages becomes,

$$P_n = P_0 \text{CTE}^{3n} \quad (4.1)$$

Rearranging eq. (4.1), gives,

$$\text{CTE} = \frac{P_n^{1/3n}}{P_0} \quad (4.2)$$

A loss of 1% of the initial charge packet in a 3-phase, 256 stage CCD requires,

$$\text{CTE} = 0.99^{1/768} = 99.99\%$$

Similar performance in a 1024 stage CCD requires  $\text{CTE} = 99.999\%$ , with the state of the art being closer to 99.9999[42] at this time. Therefore, the size of CCD imagers is fundamentally limited by the need for ever higher values of CTE.

#### 4.4.3 Power dissipation

The power dissipation in CCD image sensors is relatively high, due to the capacitance of the MOS-C gates which must be continually charged and discharged to accomplish

charge transfer. The average power,  $P_{avg}$ , dissipated in charging and discharging a capacitance,  $C_{bit}$  to a voltage  $V$  at a rate of  $f_c$  is given by,

$$P_{avg} = C_{bit} V^2 f_c \quad (4.3)$$

A typical commercially available area image sensor has 512 by 512 pixels. These are accessed through the use of vertical and horizontal CCD analog shift registers. The total capacitance to ground for each of the 3 horizontal transfer clocks is,  $C_{bit} = 75$  pF [42]. For a frame rate of 60 Hz, this capacitance must be charged up to 12 V and then down to ground at a rate of  $f_c = 16$  MHz. Therefore, the total power dissipated in driving the horizontal shift register is,

$$P_{avg} = 3 \times (75 \times 10^{-12}) \times (12)^2 \times 16 \times 10^6 = 518 \text{ mW}$$

The same sensor, when housed in a camera with associated driving electronics, is listed as dissipating 3.5 W of power from  $\pm 15$  V and  $\pm 5$  V supplies.

## 4.5 The rise of CMOS imaging

As much as CCD based imagers have advanced image sensing over the past quarter century, CMOS imaging technology is poised to revolutionize the field. The CCD will most likely remain useful in certain applications; however its future as a general purpose image sensing technology seems bleak [43] [44]. For reasons discussed below, CMOS-based imagers are the technology of choice for future implementations.

The principal trend in competitive electronics continues to be the drive towards smaller, more integrated systems. Microcontroller chips are now routinely manufactured with on-board A/D converters [45]. Telecommunications processor chips incorporate switched capacitor analog filters for post and preconditioning of signals [46]. By merging previously separate analog and digital circuitry on one chip, manufacturing costs are reduced, and noise is minimized. At the same time, because of the significant increase in system connectivity afforded by integration on one substrate, massively parallel architectures become feasible. These factors have resulted in tremendous advances in fields such as telecommunications and robotics which are highly dependent on digital/analog mixed signal circuits. Unfortunately, imaging has yet to benefit from this type of advance simply because it remains fixated on outdated

CCD technology.

When first introduced, the CCD was the only technology capable of fulfilling the goal of a solid-state television sensor [37]. Whereas successful MOSFET-scanned photodiode cameras remained elusive until the 1980's (Section 4.3), the requirements of high resolution and efficient serial readout were easily met using the CCD. By the time semiconductor manufacturing had reached the point where it was possible to integrate enough photodiodes and circuitry to make CMOS cameras viable [22], the CCD was well on its way to becoming the *de facto* standard in image sensing. At present CCD chips hold a monopoly in almost all applications ranging from medical imaging and machine vision systems, to consumer electronics. However with the introduction of mature CMOS-based sensing technology, this situation is expected to change in the near future [43].

The principal point of contention with CCD based imaging is its lack of flexibility. Whereas CMOS technology allows the integration of analog light sensing and complex digital processing on the same substrate, similar CCD implementations have remained out of reach. As was seen in Section 4.4 above, the amount of power required to drive the gates in a CCD is prohibitive. This leads to an increase in the temperature of the substrate of driving electronics, resulting in an unacceptably high level of thermally generated dark current. For this reason, it is standard practice for CCD imaging and driver electronics to be integrated on *separate* substrates in order to maintain imager performance [47]. This precludes the implementation of a completely self sufficient microchip camera.

Although significant attempts have been made at adapting the standard CCD process for general purpose applications [48], [49], the convenience associated with CMOS has not yet been matched. CMOS is a proven technology with a long history of use for both digital and analog signal processing. It is widely available through various foundries around the world and is therefore inexpensive as compared to specialized processes [46]. CCDs typically require high clock voltages (8 V - 12 V) thus ruling out their use in new and important battery powered applications such as portable computers, cellular communications, and mobile robotics. As was seen in Section 4.4, the CTE value of a CCD must be as high as possible for good sensor performance.

Increasing the size of an area or line imager requires an equivalent increase in CTE. For this reason, and due to the well known growth in number of manufacturing defects with circuit area [50], the size of CCD imagers is fundamentally limited. Finally, as was explained in Section 4.4, the very nature of charge coupling requires that data be read out in a fixed, serial fashion. This means that readout of subsets of the image data to increase system speed, or perform tracking operations, is not possible.

Therefore, it appears that the charge coupled device may soon be displaced by more flexible CMOS imaging technology. The following sections explain the basic types of CMOS imagers presently under consideration for machine vision applications. Although not strictly limited to vision-chip implementation, CMOS image sensing technology is on the verge of revolutionizing this field.

## 4.6 CMOS image sensors for computer vision

The marriage of image sensing and image processing through CMOS imaging technology is ideally suited for computer vision applications. Over the past decade numerous CMOS-based sensors for machine vision have been reported [51], [52], [53], [54], [55]. Although often very different in terms of system architecture, these sensors all benefit from the reduced power consumption, increased system integration, and ease of manufacture inherent in CMOS-based circuits. This section examines the three main types of CMOS imager and cites examples of these drawn from the literature.

### 4.6.1 The smart pixel sensor

The first type of CMOS vision sensor architecture sacrifices image resolution in favour of processing power. As shown in Fig. 4.3, the smart pixel type sensor uses a matrix of identical processor cells or *smart pixels*. Each cell is responsible for transducing the light falling in its physical vicinity, and making computations based on the value of its own photocurrent and that of its immediate neighbors. Data processing can be performed in either analog or digital mode. The results of this computation are then stored for later readout off chip.

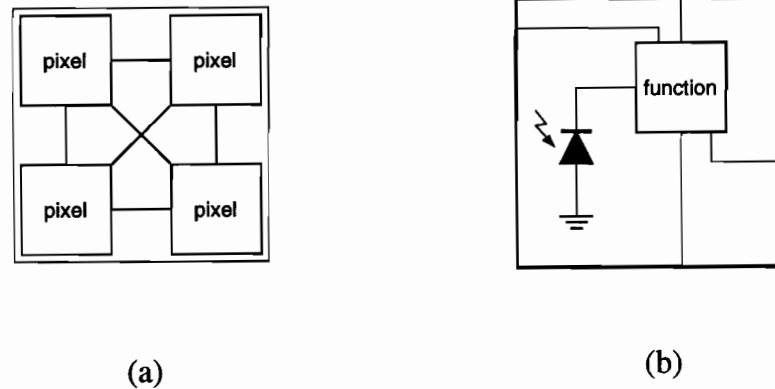


Figure 4.3: The smart pixel imager. a) Each pixel in the sensor matrix communicates with its closest neighbors. b) Each smart pixel is composed of a photosensor and associated processing circuitry

The digital smart pixel sensor makes use of simple, local binary operations to perform image processing tasks. Bernard *et al* [56] have demonstrated a “Neighborhood Combinatorial Processing (NCP) Retina” chip. Each NCP cell transfers information to its neighbor and performs boolean operations under the influence of an instruction set. Marriott *et al* [57] demonstrate a binary edge detection chip in which each smart pixel computes the thresholded spatial average of the image value at its location based on information from surrounding cells. The advantage of the digital implementation is that it is insensitive to noise; however the hardware within each smart pixel tends to be relatively complex due to the need for A/D conversion and binary storage and computation.

The analog smart pixel sensor uses nonlinear circuit elements to perform a calculation based on models of physical systems. Mead *et al* [54] demonstrate a  $48 \times 48$  pixel array which simulates the first stages of the human retina. Each processing cell takes the natural logarithm of its light input and injects the result into a resistive network which then computes a spatial smoothing function over the image. Chong *et al* [58] fabricated a sensor which determines image velocity using signals propagating

on delay lines. Each pixel contains a phototransducer and circuitry necessary for the computation. The advantage of analog implementations is the freedom to perform more sophisticated computations than with digital hardware, in a much smaller unit cell. However, the ease with which analog signals are handled is often undermined by the unavoidable noise in the system.

The success of a given smart pixel implementation hinges on the complexity of the computation being performed. Sophisticated computing machinery requires significant semiconductor real-estate to be implemented. Therefore, the smart pixel approach is limited to low resolution applications, requiring fast solutions to simple problems such as spatial filtering and image segmentation.

#### 4.6.2 The active pixel sensor

The second type of CMOS vision sensor offers increased image resolution and convenient information access, at the expense of reduced speeds of computation. As illustrated in Fig. 4.4, the *active pixel sensor* [43] is composed of an array of cells which transduce local image information. Each active pixel stores the value of its last intensity sample in an analog fashion. The data can then be accessed in a non-destructive read operation. Therefore, the array behaves as if it were an analog Read Only Memory (ROM), containing the intensity values of the last transduced image.

The advantage of such a sensor is that it provides dedicated information access to facilitate image processing. Traditional image processing systems are composed of CCD video cameras which transmit a fixed image in serial format. This image is then stored in video RAM on a host computer where the information can be accessed for subsequent processing. The link between CCD and video memory is a considerable bottleneck in the overall system and often transmits information which is eventually discarded as redundant. In the case of the active pixel sensor architecture, the sensor *is* the video RAM. The host computer is free to request exactly which information is needed, and at exactly what point in time. Therefore, as opposed to the smart pixel sensor, image processing is done *outside* of the sensing array with some loss in speed and parallelism, however with a considerable amelioration of image quality.



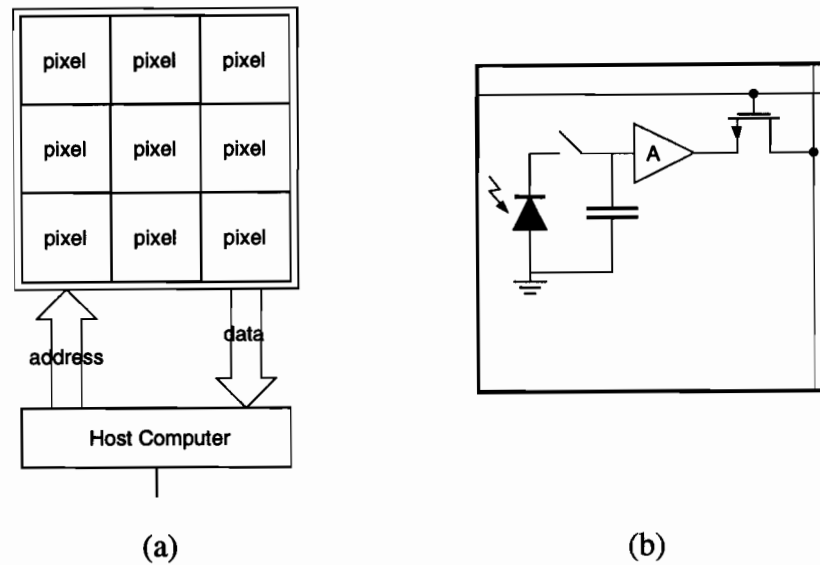


Figure 4.4: The active pixel sensor. a) The sensor behaves as a video RAM, performing image capture and storage. Data is read by the host computer through digital addressing. b) Each pixel contains a phototransducer and a buffered analog storage element. The pixel is accessed through a selection MOSFET.

The most popular active pixel sensors use photodiodes in the photon flux integration mode. As illustrated in Fig. 4.5a, the unit cell contains a reverse-biased photodiode which is periodically reset to the bias voltage through MOSFET switch,  $M_1$ . The final value of the voltage,  $V_o$ , on the photodiode after the integration period is then transferred to the storage capacitor,  $C_s$ , through MOSFET switch  $M_2$ . With  $M_2$  off, the value across  $C_s$  can be maintained for an extended period and read out nondestructively through the buffer of gain  $A$ . Yadid *et al* [59] demonstrate an  $80 \times 80$  active pixel sensor array with readout accomplished through row and column selection. Tremblay *et al* [60] developed a sensor based on a hexagonal tessellation of active pixels cells. Termed MAR for Multi-port Access photoReceptor, the device is addressed using three select lines and outputs the intensity value at the pixel of interest, as well as those in its immediate vicinity. Mendis *et al* [61] demonstrate CMOS active pixel sensors using both photodiodes and MOS-C photogates.

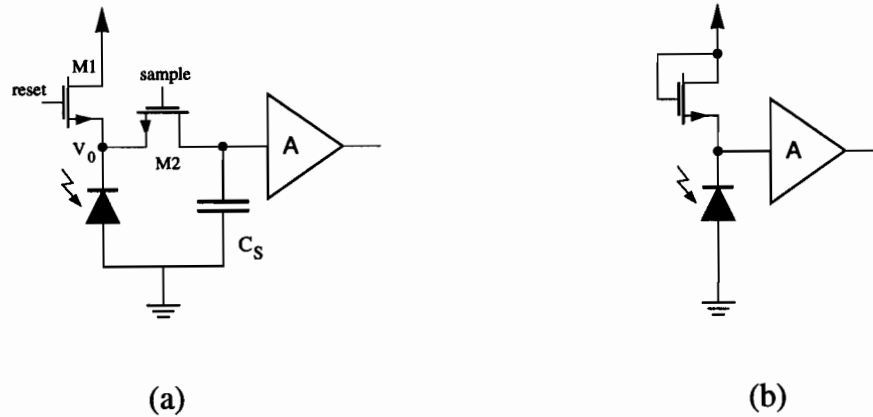


Figure 4.5: Examples of active pixel cells. a) The final value of  $V_0$  after integration is sampled and held on the storage capacitor,  $C_s$ . b) A MOSFET in the subthreshold region of operation compresses the dynamic range of the incident light signal.

An alternative form of active pixel sensor makes use of the subthreshold region of MOSFET operation to allow for instantaneous image capture. As was explained in Chapter 3, when a MOSFET is operated in the subthreshold region the source voltage is a logarithm function of the current flowing through the transistor. As shown in Fig. 4.5b, if the transistor is connected to a photodiode and its source voltage buffered, the output of the buffer amplifier will track the log of the photocurrent instantaneously. Therefore, the dynamic range of the input intensity is compressed, and access to image information is immediate since no period of integration is required. Chamberlain *et al* [29] demonstrate such a device with a light intensity dynamic range of greater than  $10^7$ . Ricquier *et al* [62] developed a  $256 \times 256$  pixel area sensor based on the logarithmic pixel concept.

#### 4.6.3 The passive pixel sensor

The final type of CMOS vision sensor provides maximum image resolution at the expense of destructive and complex data access. This sensor is a mature form of

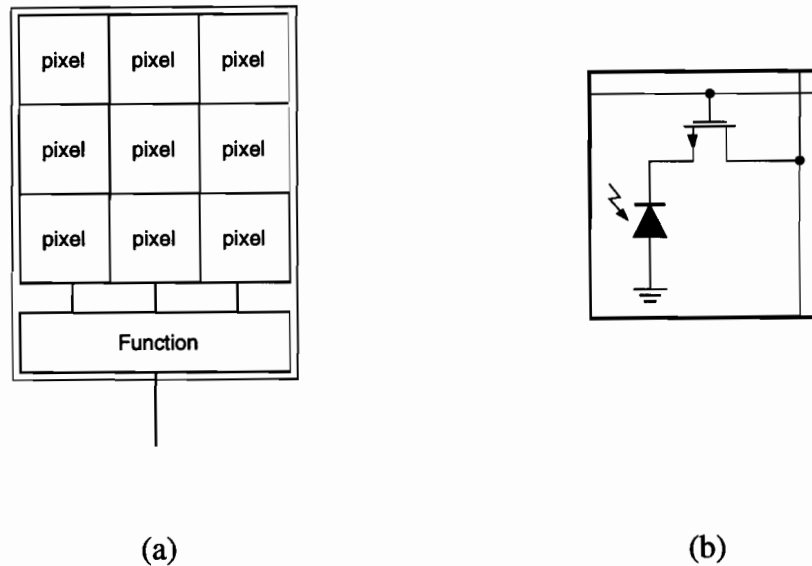


Figure 4.6: The passive pixel sensor. a) Sensor architecture for onchip parallel processing. b) Each passive cell is composed of a photodiode and a MOSFET selection transistor.

the original MOSFET-scanned diode arrays discussed in Section 4.3. As illustrated in Fig. 4.6, the sensor architecture consists of an array of passive cells. Each cell contains a photodiode and a selection MOSFET. Image capture is by photon flux integration, and readout may be in serial or parallel format. Due to the extremely small size of photodiodes in these sensors, special charge-sensing amplifiers are used to read the analog information from the cells.

The first type of passive sensor provides serial data readout. This device is essentially a modern version of the MOSFET-scanned diode array examined in Section 4.3. The serial output of this sensor makes it ideal for use as a CMOS video camera. As opposed to traditional CCD video cameras which require external drive and signal conditioning circuitry, the CMOS camera is entirely contained on one chip. Denyer *et al* have demonstrated a single chip CMOS camera fabricated in a standard  $1.5\ \mu\text{m}$  process [63]. The device has an array of  $312 \times 287$  passive cells, and draws about 200 mW from a single 5 V supply. It has automatic gain control for exposure correction,

and outputs a standard composite video signal. The same group has manufactured a complete fingerprint verification system on one chip based on this image sensor design [64]. This imaging technology is presently being commercialized by VLSI Vision Ltd. in the United Kingdom [65].

The second type of passive sensor capitalizes on the inherent connectivity of integrated circuits to bring together a high resolution photodiode array and powerful image processing machinery on one substrate. As illustrated in Fig. 4.6a the photodiode array is accessed in parallel one row at a time with each block of data input to a parallel processor at the bottom of the sensor. With this sensor architecture, such typical machine vision tasks as thresholding and convolution are performed in a very efficient manner on chip. Therefore, the traditional bottleneck of the CCD to host computer serial data link is eliminated, along with the excess mass and power consumption of those systems. Chen *et al* [51] have fabricated a sensor called PASIC (for Processor-A/D converter-Sensor Integrated Circuit) based on this architecture. The imaging array contains  $128 \times 128$  photodiodes at a pitch of  $60\mu\text{m}$ . Image data is read out in row-parallel fashion, 128 pixels at a time to 128 8-bit A/D converters. This information is then processed by 128 Processor Elements each capable of traditional Boolean functions such as AND, OR, addition and subtraction. A Swedish firm, Integrated Vision Products (IVP), has already begun marketing a similar image processing chip called MAPP [66], which has been used in applications ranging from automobile assembly lines to manufacture of high quality printing paper.

One of the major problems with this architecture is the need to fit each processor element within the pitch of the sensor pixel. This is a variation on the resolution problem inherent in the smart pixel CMOS vision sensor (Section 4.6.1). This problem might be solved through the use of sampled analog processors operating in the current mode.

## 4.7 Summary

The aim of this chapter was to provide a general overview of the history of solid-state image sensing to the present day. It began with a discussion of MOSFET-scanned

photodiode arrays which were the first semiconductor based imagers. These arrays were soon displaced by CCD based image sensors which have remained the sensor of choice for the past two decades, but may soon be replaced by a mature CMOS imaging technology. Modern day CMOS imagers benefit from reduced power consumption, and low cost due to the use of standard VLSI and ULSI production lines. More importantly however, the ability to integrate sensor arrays alongside digital and analog processing circuitry on a single substrate, lays the groundwork for the implementation of extremely powerful image processing systems. With this in mind, the next chapter presents an in-depth analysis of one of the main branches of this new technology: the CMOS PPS imager.

## Chapter 5

# The CMOS PPS Imager

### 5.1 Introduction

In this chapter, the most basic CMOS imager — the passive pixel sensor (PPS) — is examined in detail. Although not as advanced as the Active Pixel Sensor (APS), this technology is still the archetype of many CMOS imager designs. Its principal advantage is the realization of high resolution sensing due to the simplicity of the sensor matrix. As will be discussed in Chapter 7, this makes it the ideal technology for the fovea part of the CMOS foveated sensor. Simplicity in layout also makes it a good choice for the periphery part of the sensor, so that both arrays are implemented using the same circuitry.

The general topology of a CMOS PPS imager is illustrated in Fig. 5.1. It can be divided into three distinct functional units: sensor matrix, selection circuitry, and output amplifier. This chapter consists of sections devoted to the examination of each block in detail. Section 5.2 explains the basic operation of the sensor matrix and highlights important factors in the design of high resolution arrays. Section 5.3 discusses the design of digital selection circuitry necessary for readout of the PPS array. The theoretical treatment of the imager concludes in Section 5.4 with a discussion of the circuitry used for extraction and amplification of the minute signal charge contained within the sensor matrix. The three separate building blocks are then brought together in Section 5.5 where an  $8 \times 8$  pixel imager based on extracted models of real devices is simulated using Spice.

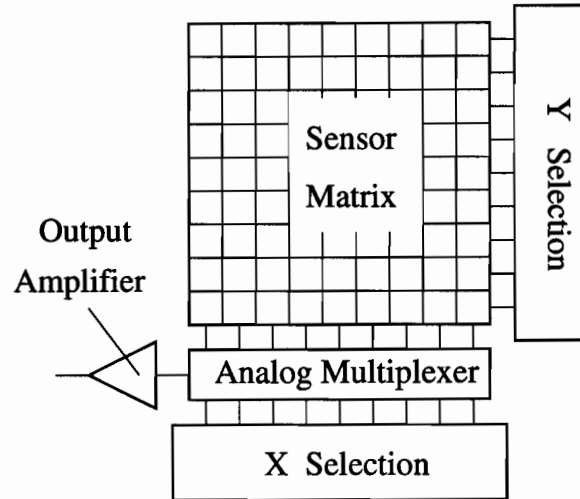


Figure 5.1: General topology of the CMOS PPS image sensor.

## 5.2 The sensor matrix

As illustrated in Fig. 5.2a, the sensor matrix consists of an array of PPS pixels. Each pixel is composed of a MOSFET selection transistor connected to a photodiode. In practise the two are usually combined, with the source diffusion of the transistor forming the photosensitive  $pn$  junction photodiode.

The drain diffusion of each MOSFET in a particular column is connected to the analog readout bus for that column. The gate of each MOSFET in a particular row is connected to the digital select line for that row. At the bottom of the array, all of the column bus lines are connected to an analog multiplexer composed of MOSFET selection transistors. Analogous to data readout in computer RAM, access to any one photodiode is accomplished through row and column select, using the digital select lines,  $Y_0, Y_1, X_0$ , and  $X_1$ .

During readout of one complete line of the image, the column select lines are each brought high successively for a *pixel time*,

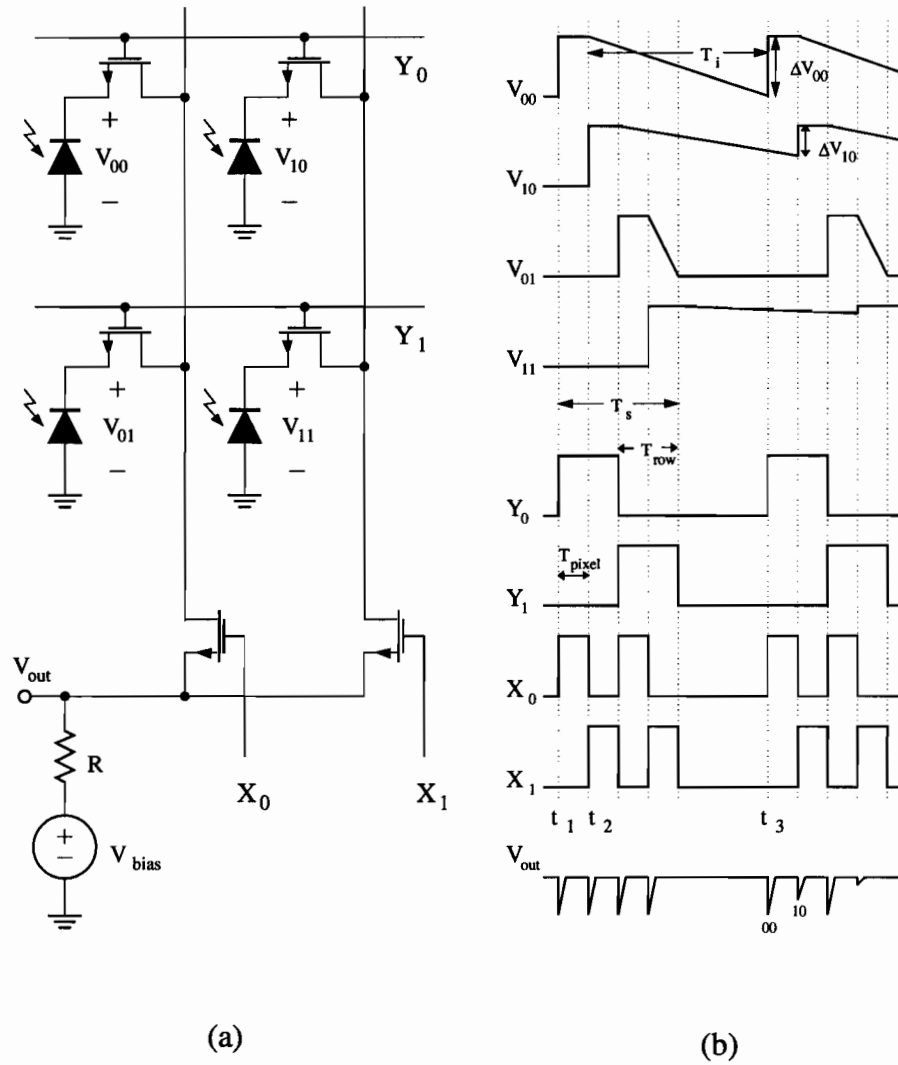


Figure 5.2: Operation of the sensor matrix. a) A 4 x 4 PPS array with row ( $Y_0, Y_1$ ) and column ( $X_0, X_1$ ), digital select lines. b) Clock waveforms for one complete frame of integration.



$$T_{pixel} = 1/f_{pixel} \quad (5.1)$$

where  $f_{pixel}$  is the *pixel clock* sometimes referred to as the *bit clock*.

Similarly, during readout of one complete frame, the row select lines are each brought high successively for one *row time*,

$$T_{row} = NT_{pixel} \quad (5.2)$$

where  $N$  is the number of pixels on one row of the imager.

The array may be operated in the photon-flux integration mode in the following manner. As was explained in Chapter 3, integration mode requires the capacitance of the photodiode to first be charged to the reference voltage,  $V_{bias}$ . This is equivalent to forcing the photodiode junction capacitance to carry the charge,

$$Q_{reset} = V_{bias} \times C_p \quad (5.3)$$

and is accomplished by selecting the particular diode in the sensor matrix, thereby charging it through the resistance  $R$  to the bias voltage  $V_{bias}$ .

As illustrated in Fig. 5.2b, each diode is selected in turn and charged to the bias voltage. Immediately after being disconnected from the output node, (for example  $V_{00}$  at  $t_2$ ), the diode voltage begins to decay due to photon flux integration. The diodes are allowed to integrate for the frame time,  $T_i$  under the condition,

$$T_i \geq T_s \quad (5.4)$$

where  $T_s$  is the time for one complete scan of the array. In most applications, image velocity is slow relative to the frame rate, and therefore the photocurrent,  $I_p$ , at each diode can be considered constant over  $T_i$ . By Eq. (3.15), the change in diode voltage  $\Delta V_{00}$  due to photon flux integration for one frame period is,

$$\Delta V_{00} = \frac{T_i}{C_p} I_p \quad (5.5)$$

Therefore, the final charge on the diode is,

$$\begin{aligned} Q_f &= Q_{reset} - C_p \Delta V_{00} \\ &= Q_{reset} - I_p T_i \end{aligned} \quad (5.6)$$

When the diode is reset to  $V_{bias}$  by connecting it to the output node, the amount of charge transferred is,

$$Q_s = Q_{reset} - Q_f \quad (5.7)$$

Substituting Eq. (5.6) into Eq. (5.7),

$$Q_s = I_p T_i \quad (5.8)$$

shows that  $Q_s$  is the light induced signal charge. Therefore, with this method of operation, diode reset and image data readout are accomplished simultaneously. In practice the time constant  $R \times C_p$  is so small (femtoseconds) that the net effect is a very short burst of current passing through  $R$ . Therefore the output voltage  $V_{out}$  consists of the reference level,  $V_{bias}$ , upon which are superimposed voltage spikes whose magnitude corresponds to the value of photocharge present on the particular photodiode being accessed. This can be seen in Fig. 5.2b in the  $V_{out}$  waveform in the form of negative going voltage spikes with magnitude corresponding to signal levels. For example  $\Delta V_{out00}$  corresponds to  $\Delta V_{00}$  and  $\Delta V_{out10}$  corresponds to  $\Delta V_{10}$ .

### 5.2.1 Proper scanning of high resolution PPS arrays

Special care must be adopted when scanning PPS arrays with very small pixels. For the simple four pixel imager illustrated in Fig. 5.2, it would at first appear that any scanning pattern works as long as the frame time,  $T_i$ , which each pixel sees remains constant. However, on closer examination, this is not the case. Selecting a particular row will cause *every* pixel in that row to be connected to their respective analog column signal lines. If these pixels are not all read before selecting the next row, their charge will be lost.

This effect is best understood by examining Fig. 5.3. Fig. 5.3a illustrates the state of a pixel and its associated analog bus line before a read operation. The photodiode

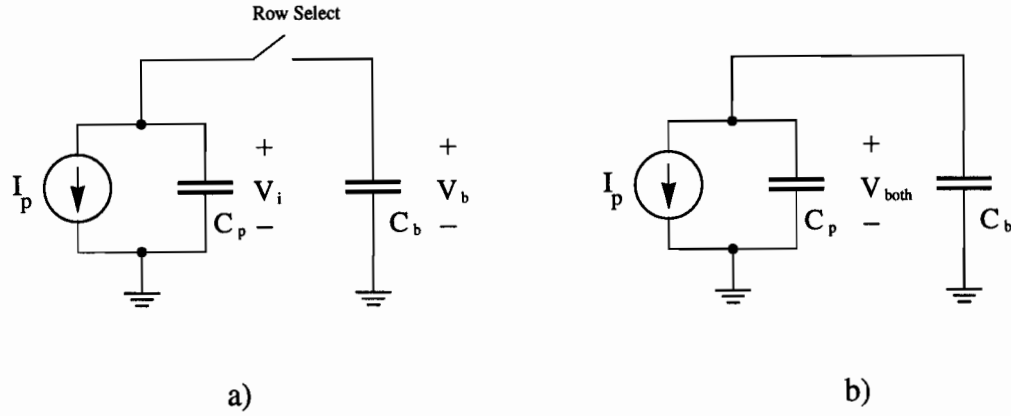


Figure 5.3: Charge redistribution during row selection. a) Prior to row selection. b) Immediately following row selection.

junction capacitance,  $C_p$  is at initial voltage,  $V_i$  and the analog bus capacitance,  $C_b$  is at initial voltage,  $V_b$ . Fig 5.3b illustrates what happens when the switch is closed. Charge redistribution causes the voltage,  $V_{both}$ , across the combined capacitance to be equal to,

$$V_{both} = \frac{C_p V_i + C_b V_b}{C_p + C_b} \quad (5.9)$$

From Eq. (5.9), it is apparent that if  $C_p$  is large compared to  $C_b$ ,  $V_{both}$  reduces to  $V_i$  and the final voltage on the pixel is essentially unaltered. Therefore, for the case of PPS arrays in which the pixels are large as compared to the signal lines, any addressing scheme will work.

However, if  $C_p$  and  $C_b$  are comparable, each time that a pixel is connected to a floating signal line, it will gain a significant amount of charge given by the  $V_b$  term in Eq. (5.9). This is equivalent to a *loss* in the signal of the pixel.

The solution to this problem is to adopt a fixed scanning procedure in which pixels are never accessed more than once in an entire frame. This is the case for instance in the imager of Fig. 5.2, if the columns are scanned quickly while the rows are scanned

slowly.

For example, the first row select line is brought high. Charge redistribution occurs between each pixel in the first row and the column signal lines. Next, each one of the column select lines are brought high in quick succession, turning on the respective column select MOSFETS. When the particular column is selected, the charge shared between  $C_b$  and  $C_p$  is read out by the external signal processing circuitry, as the column and pixel are recharged to  $V_{bias}$ . Alternately, the columns could be accessed in parallel, reducing the amount of time required to store charge on the lines. Once all of the columns have been accessed, the first row select line is brought low, and the next row select line is brought high, again forcing the pixels in that row to dump charge onto the lines.

If the row and column scanning were reversed, each time a row was selected, charge redistribution would still occur. However in this case, only the particular selected column would be read out properly. The signal charge in all the other pixels in a given row would be corrupted, and integration over an entire frame would never occur. Therefore, correct scanning is extremely important for high resolution PPS arrays.

### 5.3 Selection circuitry

As was discussed in Section 5.2, access to the information contained within the sensor matrix is accomplished by driving the address lines for vertical and horizontal selection. Various strategies exist for accomplishing this addressing on chip.

The use of decoders for both the X and Y select blocks provides the flexibility of random access to the sensor matrix. To access the information contained at a given location in the matrix, X and Y select words are fed to the respective decoders. The decoders then drive the required lines coupling the requested pixel to the output. This strategy is especially convenient in the case of APS type imagers, but is not generally used in PPS type devices.

For most applications, PPS arrays are readout in one continuous operation known

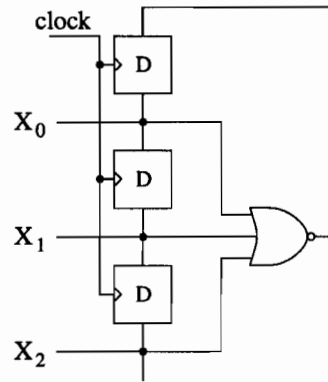


Figure 5.4: Schematic diagram of self-reset shift register.

as *progressive scan*, in which every pixel in the matrix is accessed in order. The top left pixel in the matrix is accessed first, followed by the next immediate pixel to the right, and so on until one row has been read. The next row down is then read out in the exact same way. Progressive scan continues down through the matrix until the last pixel in the last row has been read out, at which point the scan begins again from the top left pixel. The repetitive nature of progressive scan readout lends itself well to the use of digital shift registers for the X and Y selection blocks.

### 5.3.1 Self-reset digital shift register

Fig. 5.4 shows a typical self-reset shift register composed of D-type flip-flops. The outputs of the flip-flops are fed to a common NOR gate which feeds the input to the register. When all the outputs are low, a high bit is input to the first flip-flop. As the register is clocked, this bit shifts right bringing each one of the select lines  $X_0$  to  $X_2$  high successively. Finally, when  $X_3$  goes low, the output of the NOR gate goes high again to begin a new scan. Digital shift registers such as this can be implemented simply and compactly using dynamic logic, and are therefore the preferred scanning block for PPS arrays.

### 5.3.2 Shift register unit cell

The basic unit cell in the self-reset dynamic logic shift register is illustrated in Fig. 5.5a. It consists of a dynamic D flip-flop composed of a series of two inverters and two pass transistors, along with two additional transistors for accomplishing the reset operation (discussed below).

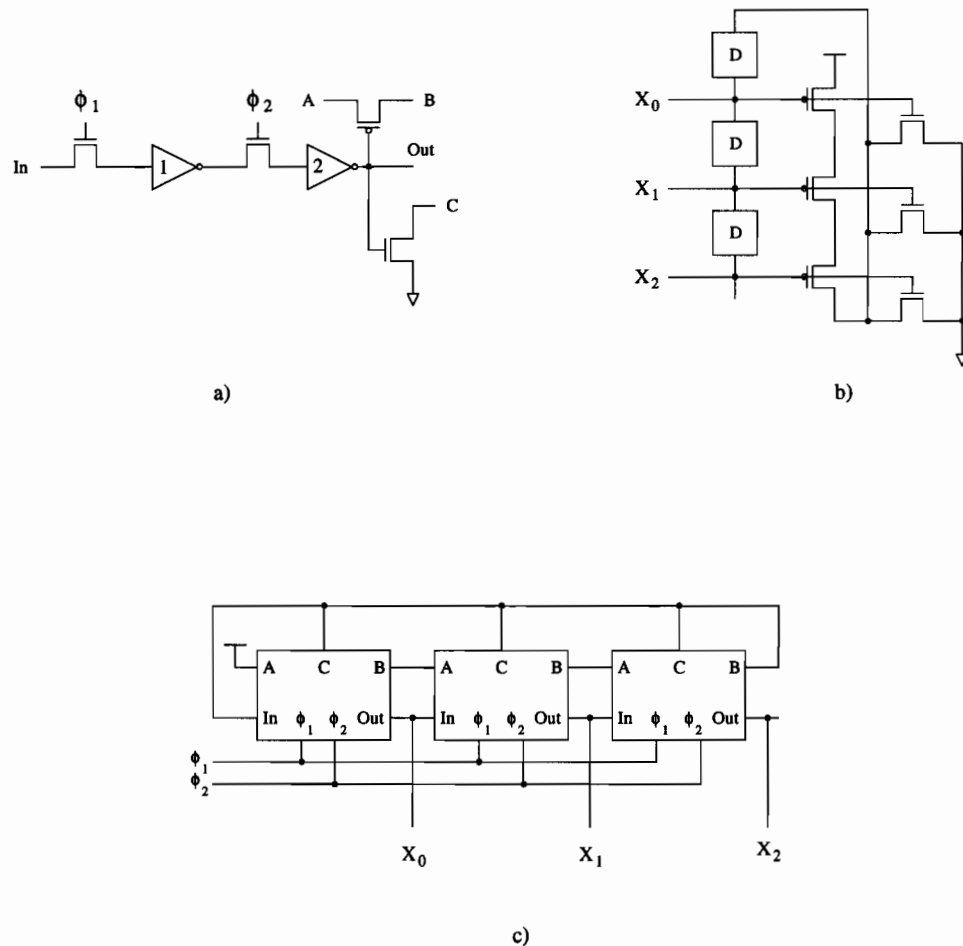


Figure 5.5: Self-reset dynamic shift register. a) Unit cell. b) Illustration of distributed NOR gate (blocks marked 'D' are D flip-flops). c) Complete 3-stage register.

The dynamic D flip-flop is operated using a two phase non-overlapping clock (discussed in more detail in Section 5.3.4) as illustrated in the figure. When  $\phi_1$  goes

high, input data is clocked to the input of the first inverter. When  $\phi_1$  goes low, this information is held on the parasitic gate capacitance of the first inverter. Next,  $\phi_2$  goes high, and the input is clocked to the second inverter resulting in output of the data from the cell. Therefore, a '1' bit can be shifted from left to right as discussed in Section 5.3.1, by alternately clocking  $\phi_1$  and  $\phi_2$ . This arrangement is convenient for very densely packed selection circuitry since it requires a minimum amount of transistors per unit cell.<sup>1</sup>

### 5.3.3 Distributed NOR gate

As was shown in Fig 5.4, a self-reset operation can be accomplished by feeding all of the outputs of the shift register to a NOR gate. To maximize the modularity of the selection circuitry, it would be convenient to *distribute* this NOR gate such that a part of it was contained inside of each unit cell. This is accomplished as shown in Fig 5.5b where each D flip-flop making up the shift register on the left is seen to have a corresponding NMOS and PMOS transistor within the NOR gate on the right. These transistors can be incorporated into the shift register unit cell as shown in Fig 5.5a, with the nodes **A**, **B** and **C** linked up among a group of cells to realize the distributed NOR operation. A 3-stage dynamic shift register with self-reset is illustrated in Fig. 5.5c, showing how input, output, and reset-related nodes are connected together to form a complete functional unit.

### 5.3.4 2 phase clock generator

As was explained in Section 5.3.2 the dynamic logic shift register requires a two phase clock for proper operation. This clock scheme is illustrated in Fig. 5.6a. Basic operation of the register requires  $\phi_1$  to be roughly the inverse of  $\phi_2$ . However for information to be clocked through effectively, there is the further constraint that the two phases must be non-overlapping as shown.

<sup>1</sup> Although in principal incident light can corrupt the operation of a dynamic logic gate by discharging nodes used for storage of information, this effect is not important at normal speeds of operation, and was not a problem in the fabricated prototypes. To achieve very low speeds of operation, it is necessary to cover sensitive circuitry with second or third layer metal.

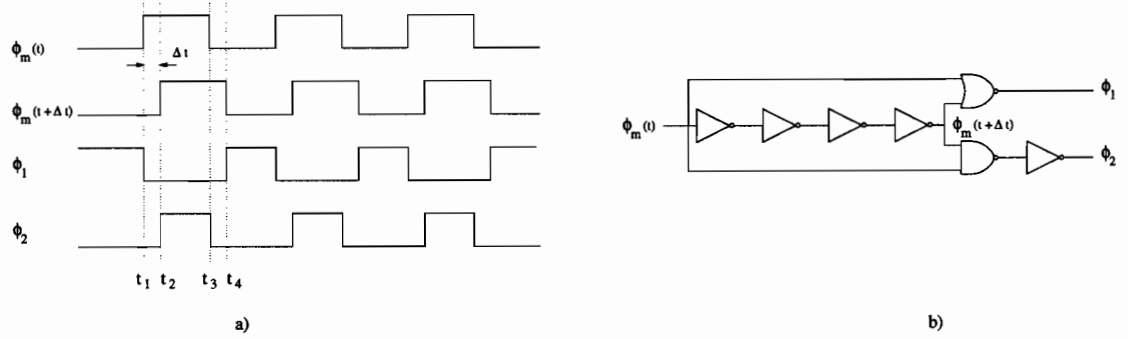


Figure 5.6: Generation of the 2 phase clock. a) Clock waveforms. b) Circuitry for clock generation.

As illustrated in Fig. 5.6a,  $\phi_1(t)$  and  $\phi_2(t)$  may be generated using a master clock  $\phi_m(t)$ , and its copy,  $\phi'_m(t) = \phi_m(t - \Delta t)$ . As can be noted in the figure,  $\phi_1(t)$  is low when either  $\phi_m(t)$  or  $\phi'_m(t)$  are high, indicating that it can be generated by the operation,

$$\phi_1(t) = (\phi_m(t) + \phi'_m(t))' \quad (5.10)$$

Similarly, it can be noted, that the second phase can be generated as follows,

$$\phi_2(t) = \phi_m(t) \cdot \phi'_m(t) \quad (5.11)$$

The logic necessary for implementing the operations described by Eq.s (5.10) and (5.11) is shown in Fig. 5.6b. It consists of an even number of inverters for generating a delayed version of the master clock, followed by the required NOR, NAND and NOT operations.

## 5.4 The output amplifier

The last functional block of the CMOS image sensor as shown in Fig. 5.1 is the output amplifier. This block is necessary to sense and amplify the small signals contained within each pixel in the sensor matrix. In this section the circuitry used to accomplish this operation is discussed.



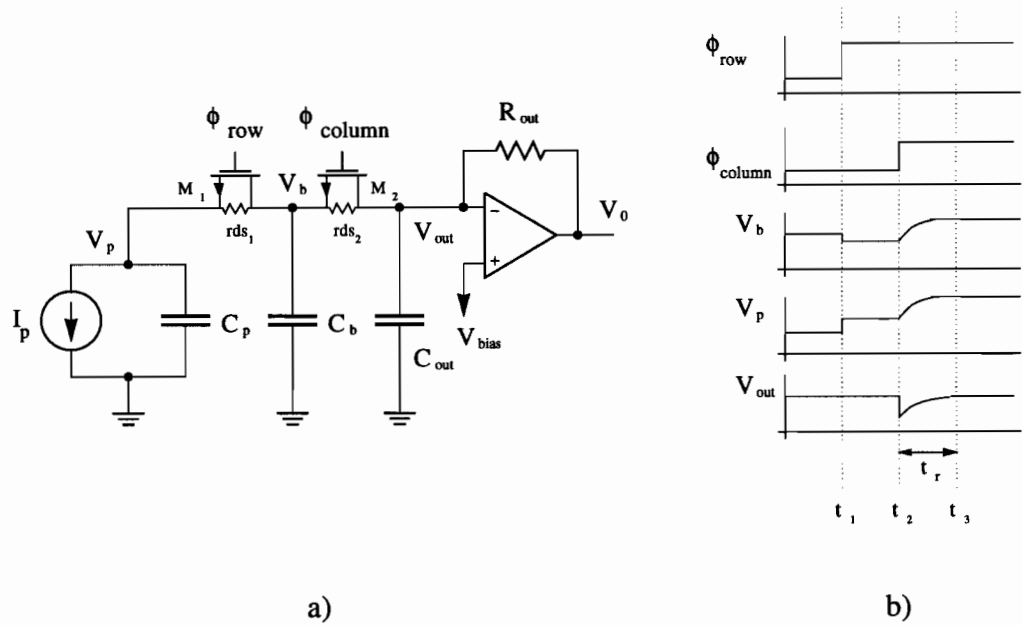


Figure 5.7: Recharge sampling. Charge is read out first to the signal line capacitance,  $C_b$ , then to the output capacitance,  $C_{out}$ . The pixel is reset to  $V_{bias}$  through  $R_{out}$ .

#### 5.4.1 Transimpedance amplifier

As was shown in Section 5.2, one convenient way of accomplishing simultaneous pixel access and reset is by placing a resistor between the output of the array and the reset voltage  $V_{bias}$ . This network is then followed by a simple voltage amplifier. The output is in the form of a series of voltage spikes whose magnitude is equivalent to the amount of signal charge at each pixel. This technique is adequate however it can be accomplished more easily through the use of a *transimpedance amplifier*. This consists of an op-amp with a resistance in its feedback path. Charge from a particular diode flows through the feedback resistance and results in a corresponding voltage on the output.

#### 5.4.2 Recharge sampling with transimpedance amplifier

Fig. 5.7a illustrates the signal path when using a transimpedance amplifier to accomplish recharge sampling. MOSFETs  $M_1$  and  $M_2$  are the row and column select

MOSFETs for the path from a pixel in the sensor matrix to the output.

The pixel is readout as indicated in Fig. 5.7b. At  $t = t_1$ ,  $M_1$  is switched, shorting  $C_p$  and  $C_b$  resulting in charge redistribution as described in Section 5.2.1. At  $t = t_2$ ,  $M_2$  is switched. Charge redistribution occurs once more, this time between  $C_p + C_b$ , and  $C_{out}$ . Once the voltage across the combined capacitance has stabilized,  $V_0$  begins to charge back to  $V_{bias}$ .

The amount of time,  $t_r$ , allotted for recharging each pixel to the voltage  $V_{bias}$  is the inverse of the pixel clock frequency,  $f_{pixel}$ . If  $t_r$  is too short, the pixel will not be completely reset to  $V_{bias}$ . This leads to *image lag*, which is characterized by the smearing of moving objects in the scene. Examining the circuit schematic of Fig. Fig. 5.7a, it would seem that the time constant of the recharge behaviour is dominated by the column bus capacitance,  $C_b$ , since  $C_{out}$  is kept very close to  $V_{bias}$  by the feedback around the amplifier. Therefore, the time constant for recharge is given approximately by,

$$\tau_r \approx r_{ds2} C_b \quad (5.12)$$

and therefore, the amount of time allotted for pixel recharge is approximately,

$$\begin{aligned} t_r &\approx 2.2 \tau_r \\ &= 2.2 r_{ds2} C_b \end{aligned} \quad (5.13)$$

As will be seen in Chapter 6, the use of nonideal operational amplifiers slows down the general operation of the circuit, and therefore, Eq. (5.13) represents the minimum recharge time for a given array.

## 5.5 Simulation example: a 64 pixel PPS imager

Illustration of the issues discussed in the previous sections is best served by examining the results of simulations performed on a simple PPS imager circuit. Although the simulator can never recreate exactly all the conditions which would exist in a fabricated prototype, the use of models extracted from a real semiconductor process<sup>2</sup>

<sup>2</sup>These models describe the same process which was used to implement the prototype of the CMOS foveated sensor discussed in chapter 7.

brings the simulation closer to reality and in some instances will point out phenomenon which might have been overlooked in the theoretical analysis.

### 5.5.1 Experimental setup

The basic circuit which was simulated is an extrapolation of the one shown in Fig. 5.2 from a  $4 \times 4$  to an  $8 \times 8$  imager. Instead of the simplified recharge method shown in that figure, the output amplifier of Fig 5.7 is used to provide gain to the signal charge.

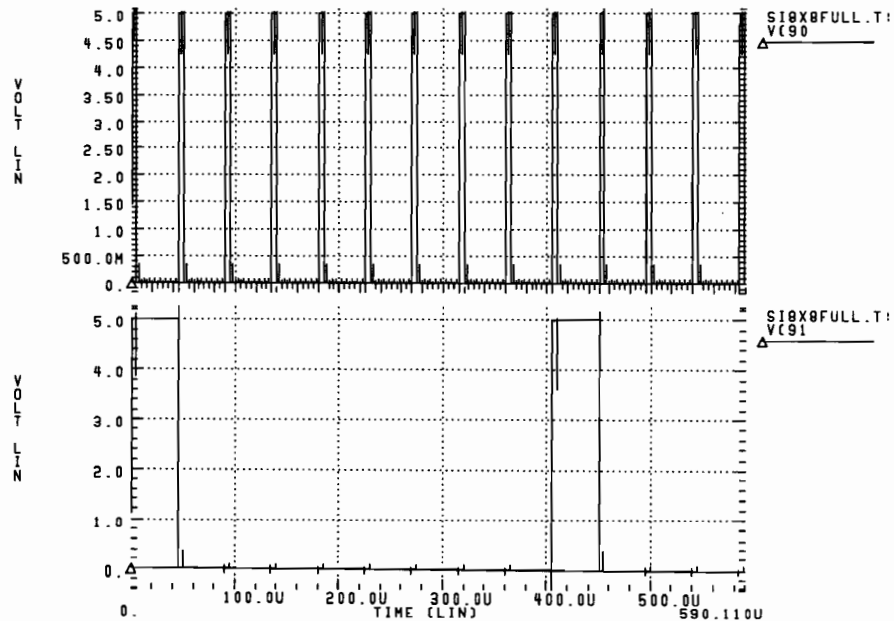


Figure 5.8: Simulation of dynamic shift registers.

The digital selection circuitry was simulated using transistors which were extracted from an actual layout of a high resolution PPS array (the fovea array discussed in Chapter 7). These are all minimum size transistors of length  $1.2\mu\text{m}$ , and width  $3.12$

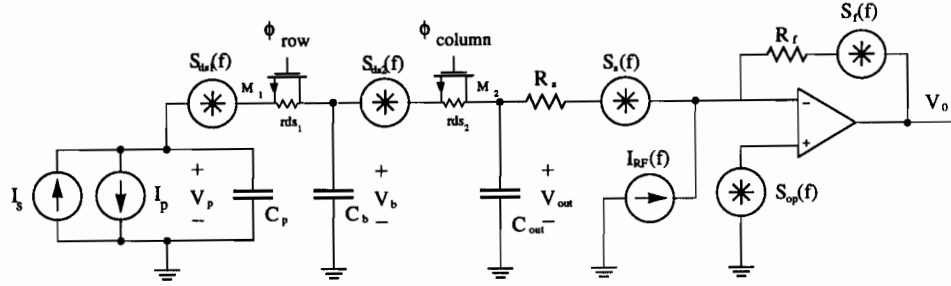


Figure 6.2: Temporal noise sources in the PPS imager.

charge. Dark current shot noise<sup>1</sup> is the limiting factor in phototransducer sensitivity at room temperature [23].

The second source of noise illustrated in Fig. 6.2 is the presence of MOSFETs in the signal path from the photodiode to the charge sensing amplifier. These devices contribute broad-band thermal noise due to their drain-source resistance  $r_{ds}$ . As indicated in Fig. 6.2 this thermal noise component can be modeled by series voltage noise sources with power spectral densities,

$$\begin{aligned} S_{ds1}(f) &= 4kTr_{ds1} \\ S_{ds2}(f) &= 4kTr_{ds2} \end{aligned} \quad (6.8)$$

where  $k$  is Boltzman's constant, and  $T$  is the absolute temperature in degrees Kelvin. These noise sources affect readout of the signal because they create random fluctuations in voltage at the output while signal charge is being read from the photodiode.

In addition to the on-chip sources described above, noise due to off-chip circuitry further affects the signal being read. Every op-amp generates a certain amount of thermal noise and this is referred to the positive input terminal as a voltage noise source with power spectral density  $S_{op}(f)$ , illustrated in Fig. 6.2. It's value depends

<sup>1</sup>It is important to note, however, that shot noise leads to uncertainty in the final value of  $V_p$  due to photon flux integration, and does not affect the signal during readout.

on the front-end stage of the amplifier, and its determination is beyond the scope of the present discussion. This voltage noise source has a flat power spectrum which is multiplied by the circuit power transfer function to yield a gained noise voltage at the output. The resistors  $R_f$  and  $R_s$  also contribute broad band thermal noise to the output and these are modeled using voltage noise sources with power spectral densities,

$$\begin{aligned} S_f(f) &= 4kTR_f \\ S_s(f) &= 4kTR_s \end{aligned} \tag{6.9}$$

as shown in Fig. 6.2.

The final source of temporal noise in the circuit of Fig 6.2 is RF pickup as modeled by the current source,  $I_{RF}(f)$ . As was indicated in Chapter 5, the minute signal charge in high resolution CMOS imagers necessitates the use of high gain charge amplifiers. Unfortunately, the transimpedance amplifier of Fig 6.2 will also amplify any variations in current at the negative input terminal due to radio waves in the environment. As will be shown in Chapter 8, it is possible to reduce this effect somewhat by shielding the high impedance node with a metal plate at ground potential.

Each of the above-described temporal noise sources contributes to a final noise spectrum  $S_n(f)$  at the output of the charge amplifier. As will be seen in Chapter 8, the output of the charge amplifier is sampled by an A/D converter for subsequent processing. This creates a sequence of noise voltage samples  $v_n(n)$  superimposed on the pixel intensity values. As will be shown in Section 6.4, the statistical variation in this data can be quantified in terms of its rms value,  $v_n$ . It is difficult to gauge the relative importance of any one of the above-described noise sources over another without a complete theoretical noise analysis. Nevertheless, as will be seen in Section 6.4 and later on in Chapter 8,  $v_n$  is readily measured and can be used to quantify the performance of a particular CMOS imager with respect to similar arrays.

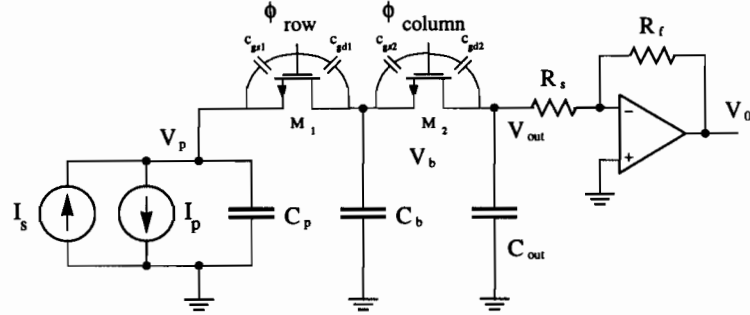


Figure 6.3: Spatial noise sources in the PPS imager.

### 6.3.2 Spatial noise

Ideally, a sensor subject to uniform illumination should produce an output image in which every pixel has the same value. Within the output image of a real image sensor however, pixel values will be different. This spatial nonuniformity in sensor output is due to random process variations in the fabrication of the device and to stray capacitance in the signal path. This nonuniformity is termed *spatial noise* or *Fixed Pattern Noise* (FPN), because it produces a distinctive ghost pattern in the image which is constant in time.

For a given device, FPN is the result of statistical variation in a number of circuit parameters. Fig. 6.3 illustrates causes of spatial noise in the signal path of a CMOS PPS imager.

The first cause of FPN is due to statistical variation in dark current  $I_s$ . As shown in Fig. 6.3,  $I_s$  competes with the signal current  $I_p$  to discharge the photodiode capacitance.  $I_s$  varies spatially across the substrate due to imperfections in the fabrication process [68]. In some cases, the magnitude of this current is so large, that the affected pixels are permanently saturated. This produces local “white spots” which degrade the output image quality to the extent that the particular image sensor may be un-

usable. Control of  $I_s$  is therefore a significant issue in improving the yeild of a given image sensor fabrication line<sup>2</sup>.

Variations in the photodiode itself directly effect the value of the signal charge which must be read out for the pixel concerned. Variations in capacitances in the signal path (as illustrated in Fig. 6.3) further contribute to FPN through the addition and removal of charge. These effects are termed *clock feed-through* and *charge injection*.

Clock feed-through is due to the parasitic gate-drain and gate-source overlap capacitances of the row and column selection MOSFETs. As illustrated in Fig. 6.3, both the column select switch, M2, and the row select switch, M1, have parasitic overlap capacitances,  $C_{gd}$  and  $C_{gs}$ .

During a read operation, M1 and then M2 will both be opened and closed in succession. When the gate voltages of M1 and M2 are brought to some high level  $V_H$ ,  $C_{gd1}$  and  $C_{gd2}$  in both MOSFETS will be charged to,

$$\begin{aligned} Q_{gd1} &= C_{gd1} (V_H - V_{bias}) \\ Q_{gd2} &= C_{gd2} (V_H - V_{bias}) \end{aligned} \quad (6.10)$$

When these switches are later opened by bringing their gate voltages to some low voltage  $V_L$ , the charge stored on the overlap capacitances will have an adverse effect on the signal being transferred. In the case of M2, charge on  $C_{gd2}$  will flow into the output amplifier, adding to the signal charge. In the case of M1, charge on  $C_{gd1}$  will combine with the charge on the column parasitic capacitance  $C_c$  and affect a subsequent read operation.

In both cases, the net effect amounts to a slight bias being added to the output signal. This would not be critical if  $C_{gd1}$  and  $C_{gd2}$  were constant over the sensor matrix. Unfortunately, due to process variations, parasitic overlap capacitances often vary and therefore contribute to FPN in the device.

---

<sup>2</sup>A study of the effects of process variations on imager defects and yeild may be found in [69].

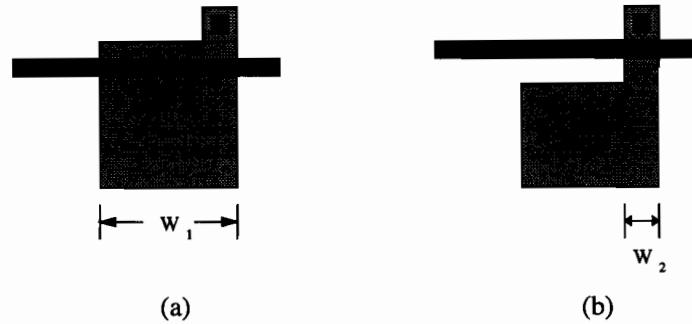


Figure 6.4: Improving pixel dynamic range through careful layout.

Parasitic capacitances also affect the starting point of photon flux integration. As with the gate-drain capacitances, the gate-source capacitances  $C_{gs1}$  and  $C_{gs2}$  are charged up when the switches are closed. When M1 is later opened, the charge on  $C_{gs1}$  will combine with the charge left on the photodiode capacitance  $C_p$  and reduce the final reset voltage left there. As was alluded to in Section 6.2 this effect is important because it limits the maximum output signal of the pixel, by reducing the starting bias voltage for photon flux integration. In order to minimize this loss, it is necessary to reduce the ratio,  $C_{gs1}/C_p$  as much as possible. The magnitude of  $C_{gs1}$  is directly proportional to the amount of overlap between the gate polysilicon and the source diffusion as well as to the area of MOSFET channel. Both of these are related to the width  $W$  of the transistor. Therefore, reducing  $W$  should decrease  $C_{gs1}$  and improve the dynamic range<sup>3</sup> of the pixel.

Fig. 6.4 illustrates how this may be accomplished. In Fig 6.4a, the selection MOSFET has width  $W_1$ . In Fig 6.4b, the selection MOSFET has width,  $W_2$ , approximately 4 times smaller than  $W_1$ . This means that  $C_{gs1} = 4C_{gs2}$ , and therefore, the loss due to clock feed-through in the pixel of Fig 6.4b is 4 times less than that in the pixel of Fig. 6.4a. Consequently, the electrical dynamic range of the second pixel is greater than that of the first. This illustrates how careful design of the pixel topology can lead to gains in phototransducer performance.

<sup>3</sup>See Section 6.4 for a discussion of dynamic range.



The net effect of each of the above-described spatial noise sources is to produce a quasi-random variation in the output image data of the device. As opposed to temporal noise, the statistical variation of spatial noise is not generally Gaussian and therefore it is more convenient to describe it in terms of the maximum variation over the entire array. Therefore, the term  $\Delta V_{FPN}$  is defined as the difference between the maximum and minimum pixel values in an image scanned from the array under zero illumination. As will be seen in Section 6.4, this leads to a useful quantification of FPN in CMOS imager arrays.

### 6.3.3 Signal processing methods for reducing noise

In general the use of differential readout is the most effective way to reduce both temporal and spatial noise in analog circuitry. With this technique, two signals are read and subtracted to produce the final output signal. The first signal is composed of the true intensity value plus random and spatial noise. The second signal is composed of only the noise. Subtracting these two yields a relatively noise free intensity output.

A number of practical applications of this concept to solid-state image sensing have been reported. Correlated Double Sampling (CDS) [70], [23] is a commonly used method by which CCD imagers are able to obtain very low values of temporal noise. This technique implements differential readout by reading the output stage of a CCD twice for each pixel of data. The first time the signal plus noise are read; the second time only the noise. The difference of the two is the final output signal. Mendis *et al* [71] have successfully migrated this technique to CMOS active pixel sensors. Their imager essentially performs differential readout within each pixel to realize excellent noise performance. To further reduce temporal noise it is possible to combine a CDS stage with an integrator or low pass filter [72]. It is also common for manufacturers of high performance line imagers to include separate dummy pixels which are physically covered in metal [40]. The signal from these pixels can be subtracted from that of the actual imaging pixels to produce a noise free output. Similar methods have been proposed for eliminating FPN on chip [38], [73].

Once the image data has been sampled and converted to digital format, a number

of digital signal processing solutions become available. As will be seen in Chapter 8, the use of a computer makes removal of FPN straight-forward: all that is required is to read a frame from the imager in the dark which forms a reference dark pattern image; this image can then be subtracted from all future images effectively removing all FPN. The removal of temporal noise is somewhat more difficult to accomplish. As will be seen in Chapter 8, it is possible to obtain excellent images from very noisy sensors if they are not used in real time; by sampling a large number of frames of the same image and averaging the value of each pixel in time, a noise free image is obtained. Depending on the quality of the original images, it may be sufficient to average as few as two or three consecutive frames for adequate results. Finally, the use of median filtering is well known as an effective method of reducing speckle noise in digital image data [10].

## 6.4 Quantifying noise

In electronics, the quality of a given signal is often expressed in terms of its *dynamic range*. This is the ratio between the maximum and minimum observable signal levels. In terms of video signals, a larger dynamic range implies more contrast, and therefore sharper, higher quality images. Therefore it is useful to characterize CMOS imagers by measuring the dynamic range of their output signals<sup>4</sup>.

As was observed in Section 6.2, the maximum signal level, or saturation output for a CMOS imager, is often set by physical limitations such as transistor cutoff and power supply levels. The minimum observable signal level is that which can be distinguished from the background noise. Therefore, to obtain the dynamic range of a given imager, it is necessary to first determine the saturation output signal  $V_{sat}$  and then measure the noise floor of the device.

As indicated in Section 6.3, a distinction is generally made between spatial noise and temporal noise in CMOS imagers. In fact, when quantifying the performance of a particular device, manufacturers of image sensing arrays provide separate figures for both types of noise examined in Section 6.3. These two noise metrics are discussed

---

<sup>4</sup>In the literature dynamic range which is the ratio of two signal densities, is sometimes confused with Signal-to-Noise-Ratio (SNR) which, strictly speaking, is the ratio of signal power to noise power.

below.

#### 6.4.1 Signal-to-Fixed-Pattern-Noise-Ratio (SFPNR)

An indication of the degree to which FPN corrupts the images produced by a given sensor is the *Signal-to-Fixed-Pattern-Noise-Ratio* (SFPNR). To measure SFPNR in a particular imager it is necessary to first obtain the dark pattern of the device. This is accomplished by averaging a large number of consecutive frames in time to remove the effects of temporal noise.

Typically, the CMOS imager of Chapter 5 produces a sequence of voltage samples  $V(n)$  corresponding to intensity values at each pixel in the sensor matrix. As the multiplexed output of the image sensor, this vector can be rearranged to form a sequence of output images. This sequence of images can be stored in a three-dimensional matrix  $I(x, y, t)$  such that the  $x$  coordinate corresponds to columns, the  $y$  coordinate corresponds to rows, and the  $t$  coordinate indexes the individual frames in time. For example, the value  $I(x_1, y_1, t_1)$  corresponds to the intensity value of the pixel at row  $y_1$ , and column  $x_1$ , which was sampled when the frame  $t_1$  was being read from the imager.

Using this formalism, the average of  $T$  consecutive frames out of a sequence of frames  $d(x, y, t)$  sampled in the dark becomes [74],

$$\overline{d_t(x, y)} = \frac{1}{T} \sum_{t=1}^T d(x, y, t) \quad (6.11)$$

And the maximum peak-to-peak dark pattern variation over the entire array is given by,

$$\Delta V_{FPN} = \max(\overline{d_t(x, y)}) - \min(\overline{d_t(x, y)}) \quad (6.12)$$

Therefore, SFPNR can be defined as the ratio between the saturation output  $V_{sat}$  and the maximum peak-to-peak dark pattern variation over the array [42],

$$\text{SFPNR} = \frac{V_{sat}}{\Delta V_{FPN}} \quad (6.13)$$

In the literature, SFPNR is usually quoted as a percentage of  $V_{sat}$ . Typical values for SFPNR range between 3.55% [71] and 0.14% [75], but in some cases can be significantly greater [76].

### 6.4.2 Dynamic range relative to temporal noise

An indication of the degree to which temporal noise corrupts the images produced by a given sensor is the *dynamic range relative to temporal noise* which is often referred to simply as the ‘dynamic range’ (DR).

To measure DR in a particular image sensor, it is necessary to first determine the dark pattern  $\overline{d_t(x, y)}$  as described above and then observe the relative variation of pixel values about this mean. This variation is roughly Gaussian in nature. For an imager with M columns and N rows, the standard deviation,  $v_n$ , for T consecutive frames can be computed as follows [74],

$$v_n = \sqrt{\frac{1}{MN} \sum_{y=1}^N \sum_{x=1}^M \frac{1}{T} \sum_{t=1}^T \left( V(x, y, t) - \overline{V_t(x, y)} \right)^2} \quad (6.14)$$

Then DR is defined as the ratio between the saturation output  $V_{sat}$  and rms pixel noise  $v_n$  [77],

$$DR = \frac{V_{sat}}{v_n} \quad (6.15)$$

Typical reported values for DR range between 72 dB [71] for high performance arrays, and 51 dB [63] for more general purpose imagers.

### 6.4.3 Performance evaluation

The quantification of noise in image sensors provides a useful measure of the relative performance of a given array in a particular application. It is important to note however, that different applications will place somewhat different constraints on performance.

For example, while arrays intended for scientific applications may display values of DR as high as 98 dB [40], for most image sensing applications a value close to 60 dB (10 bits) is sufficient. Rudimentary psychophysical experiments [74] suggest that, to an observer, variation in an image due to random noise becomes imperceptible once DR is greater than 65 dB. The same study finds that fixed pattern noise becomes imperceptible for SFPNR above 70 dB.

In computer vision applications, the requirements placed on image sensing arrays can be less stringent. While image data for these applications is typically read from CCD arrays with good dynamic range, individual pixel intensity values are stored and processed inside the computer as 8-bit numbers. The dynamic range of an image displayed on a computer screen is rarely more than 48 dB. In fact, subsequent processing of the image often reduces the dynamic range further such that the final data set consists of only binary intensity values representing various edges or interest features in the image. Furthermore, as will be seen in Chapter 8, the use of a computer makes it possible to store the dark pattern for later subtraction from image data effectively eliminating any constraint on SFPNR.

## 6.5 Charge leakage

Charge leakage occurs under intense illumination when photogenerated carriers from one pixel spread to neighboring pixels in the image. It usually arises when a region of the image contains a feature of much stronger intensity than its surroundings. For example, the bright spot denoted **A** in Fig. 6.5 leads to charge spreading in the form of a saturated region (**B**) surrounding the high intensity feature, and a bright bar (**C**) from top to bottom passing through it. These two effects are termed *blooming* and *smear*, respectively. Although both of these phenomena are caused by overillumination of the sensor, their mechanisms are slightly different and so they will be discussed separately.

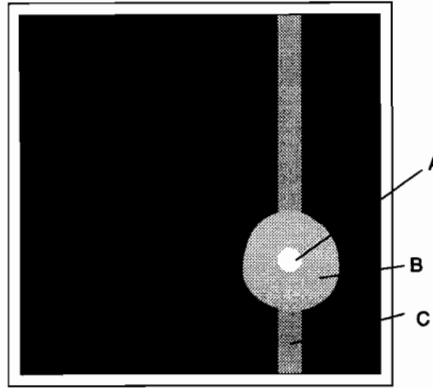


Figure 6.5: Charge leakage. A high intensity spot (A) in the image causes blooming (region B) and smear (region C).

### 6.5.1 Blooming

Blooming (region B in Fig. 6.5) is caused by charge spreading between neighboring pixels. As was stated in chapter 3, photons impinging on silicon produce electron-hole pairs. Under ordinary illumination conditions, the depletion region of the  $n^+p$  photodiode separates these charges and the voltage across the junction decays proportionately to the incident light. Under high illumination conditions, however, the photodiode becomes saturated as was discussed in Section 6.2. The depletion region of the photodiode is no longer capable of storing photogenerated charges and excess charge ‘spills over’ into neighboring pixels.

Several methods have been developed for alleviating blooming. Koike *et al* [22] propose fabricating the photodiode array within a separate well on the substrate. They form  $n^+p$  photodiodes within a  $p$ -well on an  $n$ -type substrate. The resulting  $npn$  structure is said to effectively channel excess charge away from neighboring pixels. As explained in chapter 4, CCD imagers also use charge integration and are therefore just as prone to blooming as CMOS PPS arrays. Jansson *et al* [75] have adopted the standard solution for charge spreading in CCD’s: the antiblooming diode. They place a diode-connected MOSFET within each pixel which is forward-biased when the pho-

todiode voltage drops below a certain level. This allows excess charge to be removed through a special signal line. Renshaw *et al* [53] make use of the analog signal bus along each column of pixels in their imager to evacuate excess charge. Since the drain of each MOSFET selection transistor is itself an  $n^+p$  junction, properly biasing the signal lines allows each of these to act as an antiblooming diode.

### 6.5.2 Smear

Smear is caused by charge spreading into the signal line diffusions and by light scattering [78]. Fig. 6.6 illustrates these two mechanisms of smear. Light path **A** is blocked by the metal signal line which is directly connected to and covers the drain diffusion of the selection MOSFET. Therefore this path cannot generate signal charge. However, as was explained in chapter 2, photogenerated charges within a distance  $L_n$  can diffuse into the depletion region of an  $n^+p$  junction. Therefore, as illustrated in Fig. 6.6, light path **B** will generate electron-hole pairs capable of reaching the drain diffusion of the selection MOSFET. These charges will be read out through the signal lines, even though they do not fall within the prescribed area of sensitivity of this particular pixel. The second cause of smear is illustrated by light path **C**. Here, a light wave is partially reflected at the boundary between the bulk  $\text{SiO}_2$  and the  $n^+$  diffusion. It is then inadvertently reflected by the very metal line which is shielding the drain diffusion from light path **A**. In this case, the drain diffusion is just as much a photodiode as the source diffusion.

Under high intensity lighting conditions, the above mechanisms will produce a bar across the image as illustrated in Fig. 6.5. The orientation of the bar is the key to understanding how it comes about. As was discussed in Chapter 5, high resolution PPS arrays must be scanned in two phases: first pixels dump their charge onto the capacitances of signal lines in parallel, next charge is read off of each signal line in quick succession. During the time when the signal line capacitances hold analog data waiting to be read, current generated by the two previously described mechanisms will leak away charge through the drain diffusions connected to each line. If the signal line capacitances are not read quickly enough, the amount of charge lost will be significant. In the worst case, all the charge on a given signal bus will be depleted before it can be read, producing a saturated output for the corresponding pixels. Therefore,

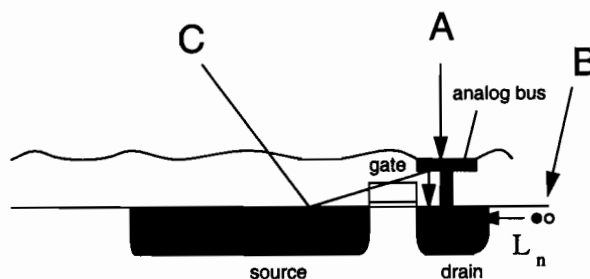


Figure 6.6: The two mechanisms of smear. Light path A is blocked by metal signal line, however, light paths B and C directly contribute to smear.

imagers in which analog signal lines run vertically with respect to the output image, will exhibit smear in the form of vertically oriented saturated bars passing through the high intensity feature in the image. Imagers in which signal lines run horizontally will show horizontal bars.

Since there is no way to prevent the signal line drain diffusions from acting as photodiodes, the only solution to smear is to reduce the amount of time charge must be stored on the signal lines. This may be accomplished in a number of ways. Renshaw *et al* [53] use a parallel access scheme in which each column signal line has its own charge sense amplifier. Therefore information is stored on the signal line capacitance only until the sense amplifiers have settled to their final value. Ando [79] describes a PPS imager in which each pixel contains an extra selection MOSFET which isolates the pixel from the signal line. Before reading a column of information, the signal lines are reset thus eliminating the effects of parasitic photodiodes. For the present version of the CMOS foveated sensor, the most simple method of antismear was adopted: to reduce the amount of time charge is stored on the signal lines, the lines (in this case columns) are scanned as quickly as possible. As will be seen in Chapter 8, this leads to acceptable antismear performance under ordinary operating conditions.



## 6.6 Summary

In this chapter, a number of non-ideal effects in the CMOS PPS image sensor have been examined. The saturation output,  $V_{sat}$ , was shown to be limited by the threshold voltage of NMOS transistors in the signal path as well as by parasitic effects. Temporal noise was found to be caused by random fluctuations in signal charge due to resistances in the signal path. Spatial noise was found to be caused by process variations and by parasitic capacitances in the row and column selection MOSFETs. Useful measures of these effects were introduced, and will be used in Chapter 8 to quantify the performance of the fabricated prototype sensor. The detrimental effects of charge leakage, namely blooming and smear were discussed at the end of the chapter. The extent of these effects in the fabricated prototype will be examined in Chapter 8.

The theory presented in the past chapters was intended to provide a background for the design of the prototype CMOS foveated sensor. The next chapter will be devoted to a detailed examination of this task.

## Chapter 7

# Design of the sensor

### 7.1 Introduction

In this chapter, the design of the CMOS foveated sensor is described. The sensor is based on the Hybrid model discussed in Chapter 2 which is optimized for implementation in VLSI. This model calls for two distinct imaging arrays: one for the fovea and one for the periphery. Although both of these are based on the CMOS PPS imager examined in Chapter 5, each one imposes unique design constraints and will therefore be discussed in separate sections. The chapter begins with a general overview of the sensor architecture in Section 7.2. This is followed by a detailed discussion of issues related to the design of the fovea array in Section 7.3. The remainder of the chapter is dedicated to the discussion of issues related to the design of the periphery array.

### 7.2 Sensor architecture

The prototype of the CMOS foveated sensor was fabricated in a commercially available analog CMOS process<sup>1</sup>. A photomicrograph of this device is shown in Fig. 7.1 and will be used to illustrate various architectural issues discussed in this Section.

As can be seen in Fig 7.1, the sensor consists of separate fovea and periphery arrays. The fovea array, highlighted at the center of the image is a high resolution PPS imager containing  $40 \times 52$  pixels each with pixel pitch of  $9.6 \mu\text{m}$ . The periphery array is a PPS imager with 64 rays and 16 rings of receptive field pixels arranged on

---

<sup>1</sup>Nortel's CMOS4S,  $1.2 \mu\text{m}$  2 metal, 2 poly process [26].

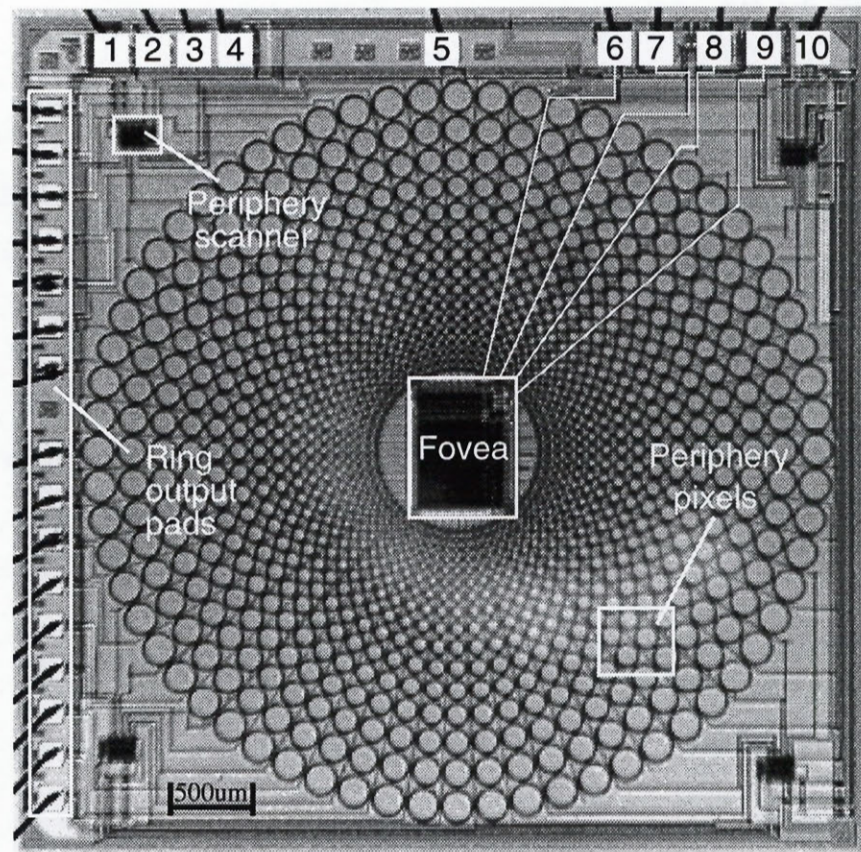


Figure 7.1: Photomicrograph of the fabricated prototype of the CMOS foveated image sensor, highlighting important features of the sensor architecture.

a log-polar grid, as was discussed in Chapter 2.

Selection circuitry discussed in Chapter 5 is integrated on-chip for both arrays. The same dynamic shift register building blocks are used for both the periphery and fovea. In the case of the fovea, these are located directly adjacent to the sensor matrix and will be discussed further in Section 7.3. In the case of the periphery, extra space available in the corners of the design is used to house the periphery scanning logic as illustrated in Fig. 7.1.

In the present implementation, the output amplifiers for both of these arrays are



located off-chip to facilitate test of the sensor components. The use of a standard CMOS process does not preclude on-chip amplification, however, and future versions of the design would benefit from improved performance with such an approach.

### 7.2.1 Readout of the periphery array

The four dark areas visible in the corners of Fig. 7.1 actually form one complete digital shift register for scanning the periphery array. These are clocked by an off-chip source through pad 4. Scanning begins near the center of the left edge of the device (at  $\theta = 180^\circ$ ), and continues counterclockwise from ray to ray until the entire periphery image has been scanned out. At this point, a periphery frame synchronization pulse is generated and taken off chip at pad 1.

All pixels on one ring of the periphery array share an analog signal bus in second layer metal. In addition, all pixels on one ray share a digital select line in first layer metal. Digital select lines are routed down in between adjacent rays of pixels along *avenues* as will be discussed further in Section 7.4. When one of these select lines is brought high, all photodiodes in receptive field pixels on the corresponding ray are connected to their respective analog bus lines. Therefore, when compared to the archetypal imager described in Chapter 5, each ring of receptive field pixels in the periphery array is analogous to one *column*, whereas, each ray is analogous to one *row*.

In contrast to the PPS array described in Chapter 5, the analog multiplexer for the periphery array in Fig 7.1 is not integrated on-chip. Instead, the rings are output in parallel via the analog pads shown highlighted along the left edge of the layout, and an off-chip multiplexer is used. The advantage of such an arrangement is that separate off-chip gain stages can be introduced for each ring *before* the analog multiplexer. This allows the gain of each ring to be adjusted independently. As will be shown in Chapter 8, this results in a properly normalized periphery output image. An additional advantage of this approach is an increase in the speed of operation of the periphery array since data is output in parallel off the chip.

### 7.2.2 Access to the fovea array

According to the Hybrid model, the fovea should be a separate array of pixels distributed on a cartesian grid in a region just inside the smallest ring of the periphery. As is visible in Fig 7.1, the fovea in the prototype sensor does not cover this prescribed region completely: it is not circular and blind spots exist on all four of its edges. As will become clear in Section 7.3, the loss of this image data is necessary in order to accommodate the digital selection logic for the fovea.

The location of the fovea array at the center of the sensor creates an interesting routing problem. A clock signal to drive the array must somehow be brought from the input at pad 9 to the selection logic of the fovea array. In addition, the multiplexed output signal from this imager must be brought from the center out to pad 6 for connection to an off-chip amplifier. Finally, power lines must be routed from pad 10 to supply the digital logic for scanning. The solution to this problem is illustrated in Fig. 7.1. Signal lines to and from the pads are routed to the fovea via unused avenues in the periphery array. Metal lines for the positive supply are distributed along 8 separate avenues in the right half plane of the sensor, providing sufficient current handling capability to drive the shift registers for scanning out image data from the center of the sensor.

The remainder of this chapter will be devoted to more detailed discussions of the individual arrays mentioned above. Important issues in layout of these two sections of the CMOS foveated sensor will be examined, beginning with a discussion of the fovea array in Section 7.3.

## 7.3 The fovea

This section details the layout of the fovea by highlighting various elements in the array topology. It then goes on to discuss the implementation of the fovea pixel cell and concludes with a brief preview of the expected performance of the fabricated prototype.



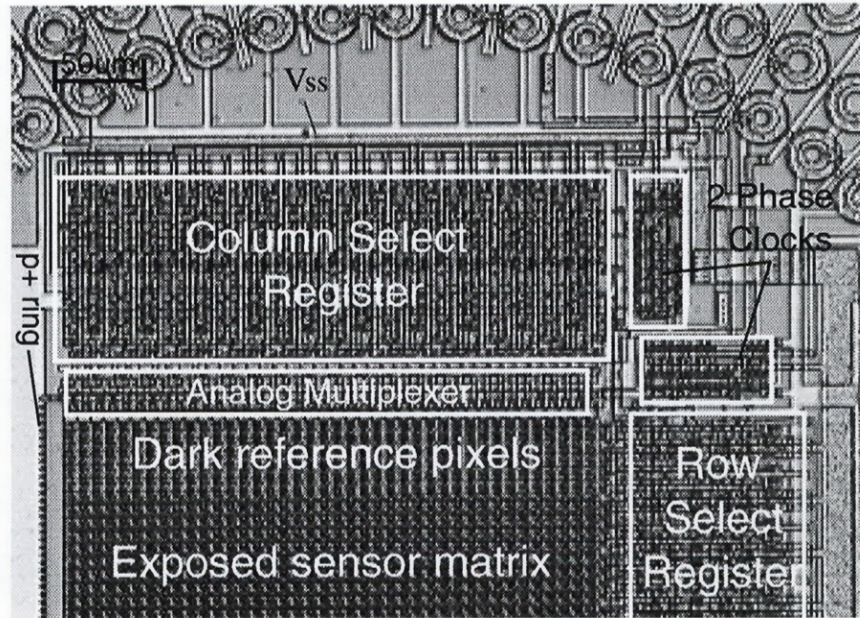


Figure 7.2: Photomicrograph of upper half of fovea array.

### 7.3.1 Fovea overview

The photomicrograph in Fig. 7.2 shows a close-up of the top section of the fovea array in which the sensor matrix, analog multiplexer and digital shift registers are visible.

#### Sensor matrix

The sensor matrix is visible in the lower half of Fig. 7.2. As explained in Section 7.2, the pixel pitch in the fovea is  $9.6 \mu\text{m}$ . This is comparable to that of today's megapixel CCD area imagers. To give an idea of the resolution which this imager obtains, covering the entire foveated sensor of Fig 7.1 with pixels this size would result in an array of  $500 \times 500$  pixels in an area of silicon only  $4.8 \text{ mm} \times 4.8 \text{ mm}$  square. To realize such a high degree of resolution, it was necessary to move substrate biasing connections outside of the sensor matrix itself. Biasing is accomplished with a ring of  $p^+$  diffusion encircling the sensor matrix kept at  $V_{ss}$  potential. This ring is partially visible on the left side of the figure.

As was discussed in Chapter 5, the presence of dark reference pixels in an imager facilitates the test of the device. The fovea array of the prototype sensor was designed with this in mind, and dark reference pixels are realized by covering 6 rows of pixels at the top, and 6 rows at the bottom of the sensor matrix with second layer metal. The top 6 rows are visible in Fig. 7.2. Ignoring dark reference pixels, the fovea size is  $40 \times 40$  pixels, and therefore most example images in Chapter 8 will be cropped to this size.

### **Analog multiplexer**

At the top of the sensor matrix in Fig 7.2, the analog multiplexer is visible. This is formed out of a series of 40 NMOS transistors whose source diffusions are connected to respective columns and drain diffusions are all tied to the output line. The gates of these MOSFETs are each connected to one stage in the column select shift register. Data is multiplexed to the output line by clocking the shift register such that it selects each one of the MOSFETs in turn.

### **Selection circuitry**

The row and column shift registers shown in Fig. 7.2 are each realized using dynamic logic as was described in Chapter 5. Each shift register has its own 2 phase clock generator to avoid any race conditions which might occur by routing multiple clock lines from a single source. The 2 phase clock generator for the column select register is driven directly by an input clock as discussed in Section 7.2.2. This signal is routed down through a periphery avenue as was described in Section 7.2. The row select register is driven by the synchronization pulse from the column select register.

### **Supply connections**

As was described in Section 7.2 positive power supply connections are routed down through avenues in the periphery array. As can be seen in Fig. 7.2, negative supply connections are taken directly from the periphery array and distributed one line per periphery ray in order to provide for proper current handling capability on this supply.



### Blind spots

It is clear from Fig. 7.2 that significant blind spots exist at the interface between the fovea and periphery arrays. The diameter of the disc contained inside the innermost ring of the periphery is roughly  $900\text{ }\mu\text{m}$ , and would amount to a fovea with a diameter of close to 100 pixels. As compared to the actual fovea size of  $40 \times 52$  pixels, it is clear that a considerable amount of image data is lost. The only solution to this problem would be to remove the scanning circuitry from the center of the sensor and place it outside the periphery array. This would require a very large number of select lines to be routed down through available periphery avenues to drive the fovea sensor matrix. Alternately it might be possible to share select lines with the periphery array. In either case the interface between the avenues which are on a polar grid, and the fovea matrix on a cartesian grid is a nontrivial problem. In the present implementation, it was decided that the best compromise would be to integrate the selection circuitry for the fovea array directly adjacent to the sensor matrix.

#### 7.3.2 Implementation of the fovea pixel cell

The maximum spatial resolution which can be obtained in a CMOS imager will always be that obtained with a PPS array. This is because, each pixel in a PPS array contains only 1 MOSFET, whereas other types of pixel such as the APS contain many more. Therefore the need for maximum resolution dictates the use of PPS technology in the fovea array.

As explained in Chapter 3, the basic PPS cell consists of a MOSFET selection transistor whose source-substrate diffusion functions as a photodiode. Fig. 7.3a shows how such a PPS cell can be formed in a standard CMOS n-well process. The source diffusion of the selection MOSFET is elongated in order to increase its sensitivity to incident light. The  $n^+p$  junction transduces incident light through photon flux integration as was described in Chapter 3. In Fig. 7.3a, it is apparent that the junction consists of the bottom plane of the  $n^+$  diffusion (region **A**), as well its sides (region **B**). Although most of the photoelectric effect takes place on the bottom plate, some charge is accumulated along the sides as well. Photons falling outside region **A** but



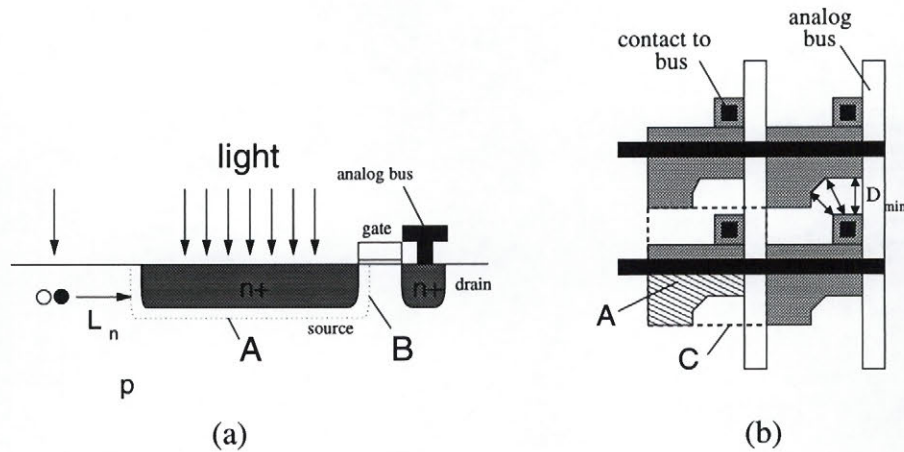


Figure 7.3: Implementation of the PPS cell in standard CMOS. a) Cross section of the unit cell. b) Plan view of the unit cell.

within the diffusion length of electrons  $L_n$  also contribute to the final signal. Therefore, to some extent, the region of sensitivity of each pixel is not entirely defined by the diffusion photodiode, and some blurring of the image takes place.

As illustrated in Fig. 7.3b, each pixel has an area  $C$  equal to the square of the pixel pitch. The smallest pixel size which can be achieved in a given process is limited by  $D_{min}$  the minimum spacing allowed between adjacent diffusion regions.  $D_{min}$  also limits the maximum area  $A$  allowed for the photodiode diffusion so that it is somewhat smaller than  $C$ . The percentage of the entire pixel area which is covered by the photodiode is termed the *fill factor* and is a measure of the sensitivity of the pixel. As is shown in Fig. 7.3b, it is possible to increase the pixel fill factor by elongating  $A$  slightly while still conforming to the design rules. Even with this elongation however, pixels in the fovea of the prototype sensor have a fill factor of about 21%. This indicates that most of the pixel area is taken up by the selection MOSFET and contact to the analog bus.

### 7.3.3 Performance issues

As was discussed in Chapter 6, various sources of noise exist in CMOS imagers. Unfortunately, as a high resolution matrix of PPS cells, the fovea array is expected to be subject to these problems.

The use of a single bias ring surrounding the entire sensor matrix means there will be inadequate drainage of excess charge for individual devices. This will result in blooming along the signal lines, as well as charge spreading from pixel to pixel. The solution to this problem would be to place bias contacts within each pixel, however this would result in reduced image resolution and therefore was not implemented.

The use of minimum sized transistors and photodetectors within the fovea array should result in increased fixed pattern noise due to process variations. In addition, the very low junction capacitance of the photodiodes leads to a saturation charge on the order of 0.1 pC. This minute signal is expected to result in very low dynamic range with respect to temporal noise sources.

As will be seen in Chapter 8, the above mentioned effects while severe, can in fact be reduced to a certain extent through the use of subsequent DSP processing. Future versions of the sensor would have to institute on-chip measures such as the use of APS pixels and differential readout to achieve optimal performance.

## 7.4 The periphery

A close-up of a region of the periphery array in the prototype sensor is shown in Fig. 7.4. This figure as well as Fig. 7.1 in Section 7.2 illustrate the highly non-standard layout geometries used in the design of the periphery array. The complexity and sheer magnitude of the implementation of such an array suggest the incorporation of some degree of automation into the process. Fortunately, as will be shown in subsequent sections, the highly symmetric geometry of the log-polar mapping lends itself readily to computerized layout.

This section describes various key issues in the design of the periphery array with



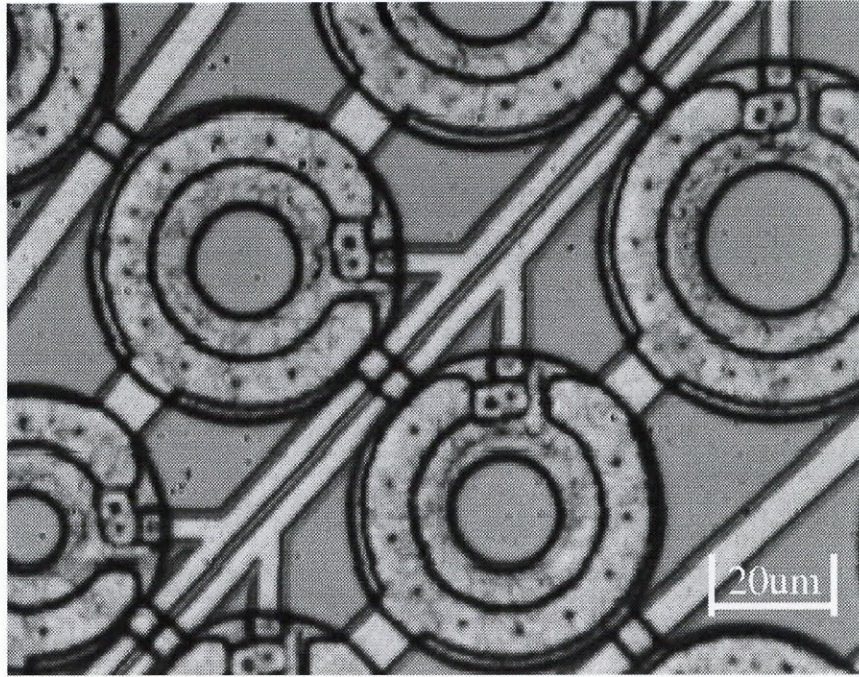


Figure 7.4: Closeup of periphery array showing receptive field pixels.

an emphasis on the development of special software for design automation. The software, termed **PGEN** (for Periphery Generator), consists of a library of C language routines for creating nonstandard geometries in Caltech Intermediate Form (CIF). The main focus of this section therefore, is the development of rules to be used by such a program for efficient and accurate generation of the periphery array.

#### 7.4.1 Periphery array topology

According to the Hybrid model developed in Chapter 2, the periphery array is composed of a unique tessellation of receptive field pixels located at the intersection of  $M$  rays and  $N$  rings. In the prototype sensor,  $M=64$ , and  $N=16$ , therefore, the spacing between adjacent rays is given by Eq. (2.17) as,

$$\Delta\theta = \frac{2\pi}{M}$$

$$= \frac{2\pi}{64} \quad (7.1)$$

The radius  $r_n$  of the  $n$ th ring of pixels in the array is calculated using Eq. (2.16) repeated here,

$$r_n = \alpha e^{n/\beta} \quad (7.2)$$

The diameter of individual pixels located on this ring may be determined using Eq. (2.18),

$$D_n = \Delta\theta r_n \quad (7.3)$$

The above equations, combined with Eq. (2.11) can be used by automated layout software to correctly size and place receptive field pixels in the periphery array. Before this can be accomplished, however, the program must solve for the constants  $\alpha$ , and  $\beta$  using constraints at the array's inner and outer boundaries.

The inner boundary of the array is located at the edge of the innermost ring ( $n = 0$ ), for which,

$$\begin{aligned} r_0 &= \alpha e^{0/\beta} \\ &= \alpha \end{aligned} \quad (7.4)$$

The diameter of cells on this ring is by Eq. (2.18),

$$\begin{aligned} D_0 &= (\Delta\theta r_0) \\ &= \Delta\theta \alpha \end{aligned} \quad (7.5)$$

Therefore, the innermost boundary of the periphery array,  $R_{in}$  is located at,

$$R_{in} = \alpha - \alpha\Delta\theta/2 \quad (7.6)$$

Solving for  $\alpha$  yields,

$$\alpha = \frac{R_{in}}{1 - \Delta\theta/2} \quad (7.7)$$

For the prototype sensor shown in Fig. 7.1, the value of the inner radius was specified at  $R_{in} = 470 \mu\text{m}$ , therefore, in this case,  $\alpha = 494 \mu\text{m}$ .

Table 8.1: Response of test chip photodiodes. Response is normalized with respect to diode No. 1.

Diode No.	Response
1	1.00
2	0.32
3	0.77

this leads to a moderate resistance to blooming at typical frame rates of operation.

The results reported above show that the chosen fabrication process possesses good optoelectronic characteristics for use as a real time imager in the visible range. The remainder of the chapter is devoted to a discussion of the test of the prototype foveated sensor which was realised through this same fabrication process.

### 8.3 Image sensing test apparatus

The characterization of the CMOS foveated image sensor required the development of special test apparatus. The complete test system includes electrical, mechanical and optical sections which for convenience were built from off-the-shelf components. Before proceeding to a discussion of the test results, it is useful to examine this experimental setup.

#### 8.3.1 Opto-mechanics

In order to characterize the sensor's ability to capture images, the test setup shown in Fig. 8.3 was developed. The device under test (D.U.T) is mounted on its own circuit board along with necessary local circuitry described in Section 8.3.3. This board is secured to a C-type lens mount using Spindler and Hoyer opto-mechanical components as shown. The lens mount holds a standard TV camera lens which focuses images directly onto the surface of the die.

Although the sensor is capable of imaging general scenes under normal lighting conditions, most of the images in this thesis were taken in a controlled environment.

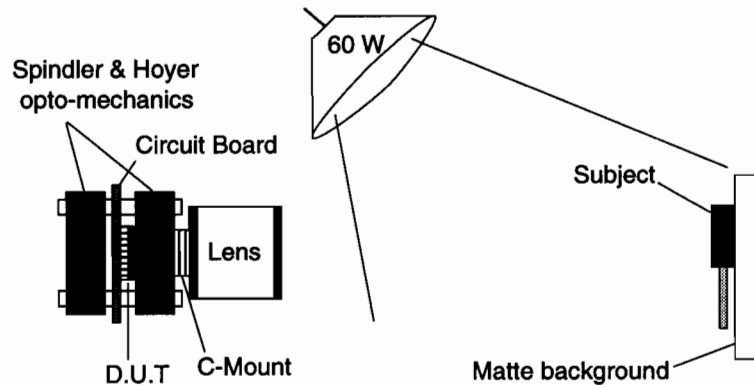


Figure 8.3: Opto-mechanical test setup for foveated sensor.

As illustrated in Fig. 8.3, this consisted of a diffuse white light source (60 W tungsten bulb), shone on a matte white background at a distance of 60 cm from the camera. Objects or laser printed patterns were placed against this background and the lens focus and aperture were adjusted until clear output images were produced.

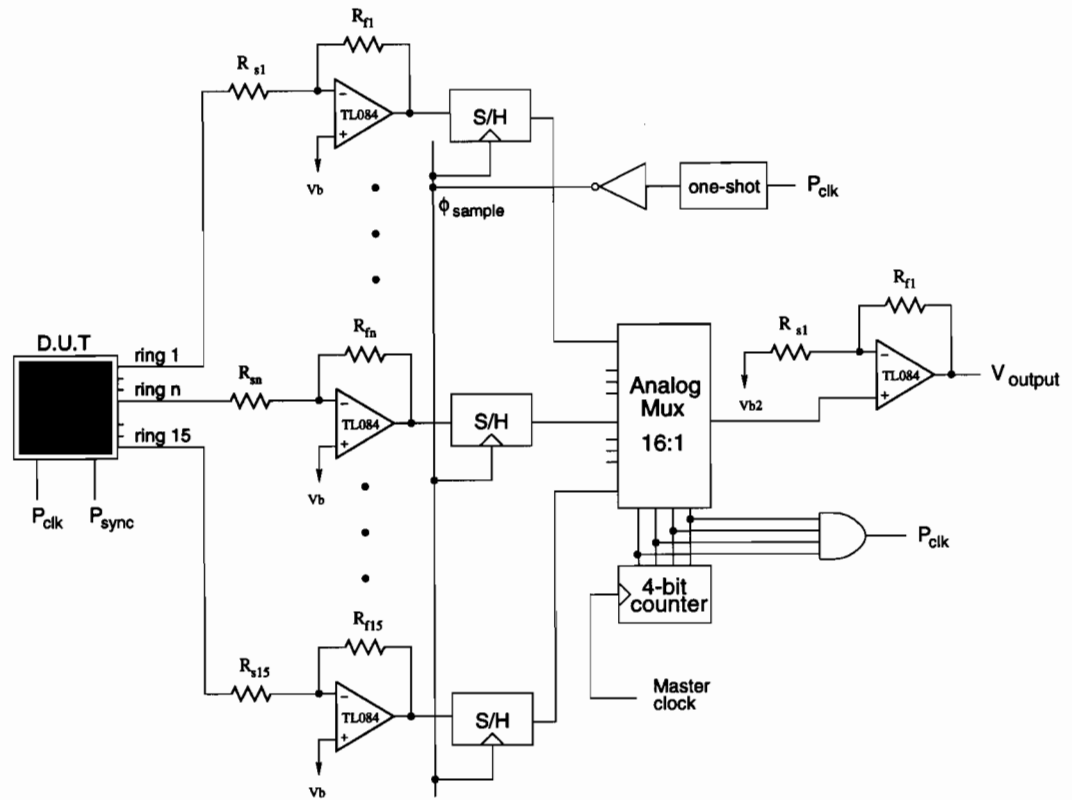
### 8.3.2 Periphery test circuitry

To test the operation of the periphery array, it was necessary to create support circuitry off-chip. This circuitry generates proper driving signals for the device, and reads and formats the analog image data so that it can be input to a computer as described in Section 8.3.4 below.

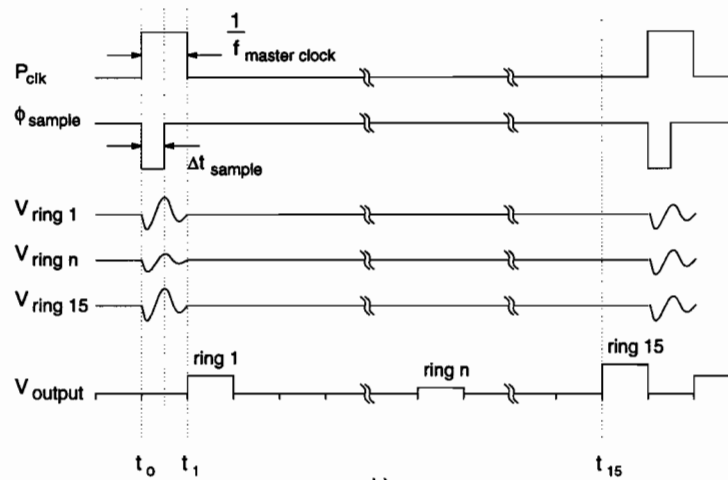
A schematic diagram of this circuitry and waveforms illustrating its operation are presented in Fig. 8.4. The circuit consists of 15 analog signal processing trains; one for each of the 15 active on-chip rings<sup>5</sup>. The outputs of these processors are time multiplexed using an analog multiplexer to produce a serial stream of voltage levels corresponding to the intensity value at each pixel in the periphery array.

As illustrated in Fig. 8.4a, each one of the 15 parallel signal processing trains is

<sup>5</sup>Note that as described in Appendix B, the fabricated prototype only outputs 15 rings of useful information.



a)



b)

Figure 8.4: Off-chip circuitry for amplification and formatting of output from the periphery array. a) Circuit schematic diagram. b) Waveforms for readout of one complete ray of information.

composed of a transimpedance amplifier followed by a sample and hold circuit. For improved performance, these could be followed by a charge integrator and a DC restorer implementing correlated double sampling (described in Chapter 6). As will be shown in Section 8.4, however, the performance with the present circuitry is adequate for general purpose applications.

The transimpedance amplifiers implement recharge sampling, and therefore provide the dual functions of pixel reset and signal amplification described in Chapter 5. The values of the resistors  $R_{sn}$  and  $R_{fn}$  are chosen to normalize the output for each ring such that ring 1 has the most gain, and ring 15 the least. This accounts for the fact that, given uniform illumination of the periphery array, rings with smaller photodiodes produce proportionately less output charge than those with larger ones. As will be seen in Section 8.4.1, this is a somewhat rudimentary scaling technique, and further digital processing of the image is required for proper output. Future prototypes using on-chip charge integrating amplifiers would be unaffected by this problem.

As was described in Chapter 5 the output of the transimpedance amplifier for one pixel of image data is in the form of a somewhat brief, damped oscillation. The magnitude of each peak in this waveform corresponds to the intensity value at the given pixel. To facilitate later processing, it is necessary to capture and store the magnitude of the first peak (the largest one) and this is the purpose of the sample and hold circuitry in the signal train in Fig. 8.4a. These 15 S/H blocks combined may be thought of as an analog storage register which at any given time during a frame holds normalized data values for the ray which is presently being accessed.

Fig. 8.5 illustrates sampling of the transient. The sample pulse is produced by inverting the row synchronization pulse ( $P_{clk}$  in Fig. 8.4b), and feeding it to a one-shot as shown in Fig. 8.4a. The one-shot delays the sample edge by an interval  $\Delta t_{sample}$ , such that sampling occurs near the peak of the transient waveform.

Once all of the pixel values for one ray have been sampled and held as described above, they are ready to be multiplexed to a single output stream. Fig. 8.4b illustrates the operation of the circuit during multiplexed readout of one ray of the periphery. At the beginning of the period at time  $t_0$ ,  $P_{clk}$  goes high. This causes the on-chip



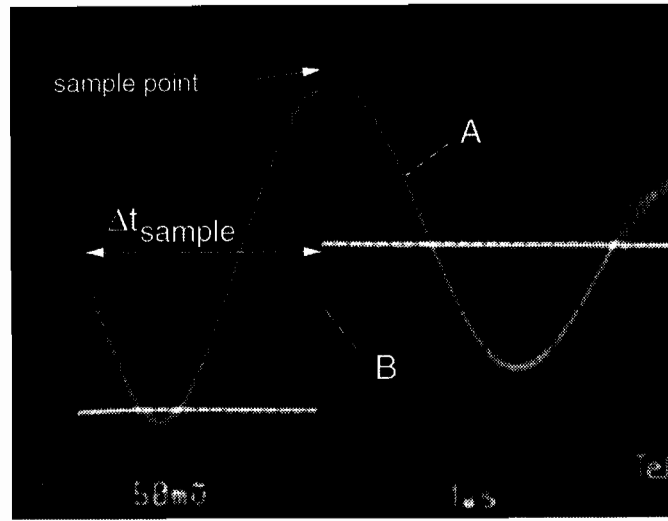


Figure 8.5: Oscilloscope trace illustrating sampling of transient at the output of the transimpedance amplifier. Signal A is the output of the transimpedance amplifier while signal B is the sample pulse,  $\phi_{sample}$  generated by the one-shot.

select logic of the foveated sensor to select the next ray of pixels for readout. Charge flows on each of the 15 analog output lines due to recharge sampling, and the off-chip transimpedance amplifiers produce transient waveforms as a result. After a period of time  $\Delta t_{sample}$ ,  $\phi_{sample}$  goes high causing the value of each of the 15 transient waveforms at that instant to be captured into the analog register as described above. At the next tick of the master clock, at time  $t_1$ , the analog multiplexer begins selecting each one of the 15 values in the analog register in sequence, ending with the last one at time  $t_{15}$ <sup>6</sup>. The output of the multiplexer is then amplified once again, and given a variable offset to simplify interfacing it with subsequent processing circuitry.

Fig 8.6 shows the multiplexed output over one ray time under zero illumination. It is important to note that significant differences exist in the dark values of the various rings in the ray. These offsets are corrected using the calibration procedure described in Section 8.4.1.

<sup>6</sup>Note that the lack of useful information on ring 16 means the off-chip support circuitry can be simplified since the period of time allotted to ring 15,  $t_{15}$ , can actually be used to generate the  $P_{clk}$  signal.

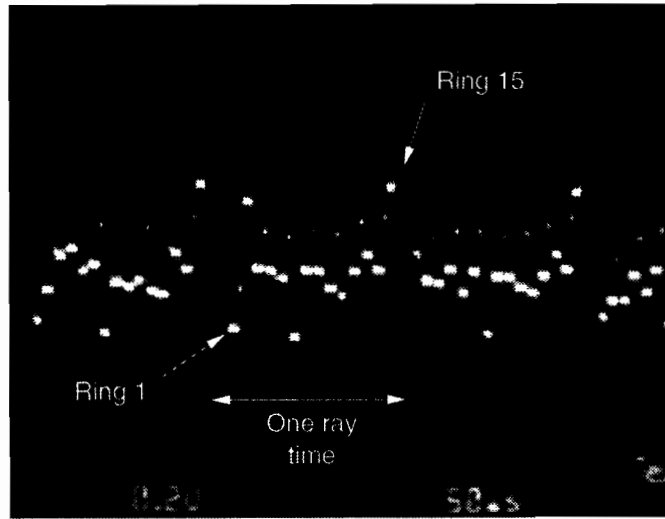


Figure 8.6: Closeup of oscilloscope trace showing periphery output waveform for 3 consecutive rays of pixels with 15 pixels each.

In order to read out one complete frame of information, all 64 rays of the periphery must be accessed. For a frame rate of  $f_{frame}$ , the  $P_{clk}$  input must be clocked at a frequency,

$$f_{P_{clk}} = f_{frame} \times (64 + 1) \quad (8.2)$$

Clocking the  $P_{clk}$  line causes the on-chip shift registers to select each ray of the device in series starting from ray 1 until ray 64. Data from each ray must be read within the period  $1/f_{P_{clk}}$  or it will be lost. After ray 64 is selected, the on-chip scanners reset automatically producing a frame synchronization pulse of duration  $1/f_{P_{clk}}$  (accounting for the 1 in Eq. (8.2)). This results in a 'phantom ray' of image data which shows up as a black line in the output image as will be seen in Section 8.4.3 below.

Fig. 8.7 illustrates the output waveform for one complete frame. In this case, the camera was pointed directly at a light bulb which shows up as a large bump near the beginning of the waveform. As will be shown in Section 8.4, the waveform of Fig. 8.7 is converted to an output image by decoding the multiplexed data in a digital computer and displaying these intensity values on a video screen.

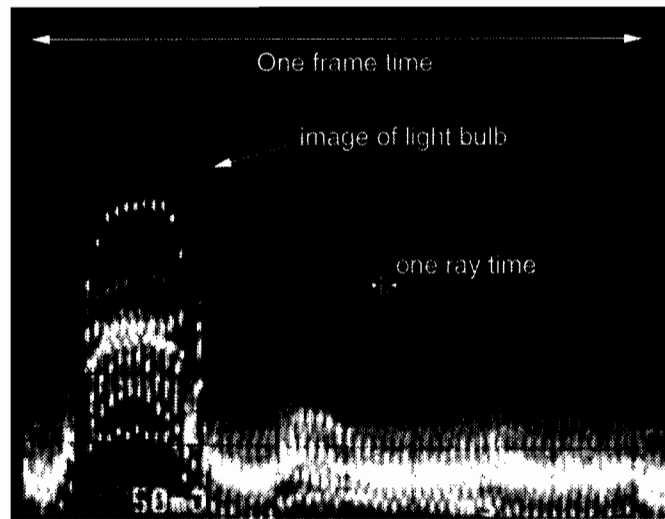


Figure 8.7: Oscilloscope trace of one complete frame of the output waveform from the periphery array. Large bump near the beginning of the frame corresponds to a light bulb in the output image.

### 8.3.3 Fovea test circuit

The test of the fovea part of the sensor required less complicated support circuitry than that described above for the periphery. The fovea contains both row *and* column scanners on-chip, and therefore the output signal is already formatted as a single stream of image data. In contrast to the periphery, however, individual photodiodes in the fovea are quite small and therefore generate very little signal charge. Consequently, the principal consideration for readout of the fovea array is optimization of signal gain.

The off-chip circuitry for the fovea array is illustrated in Fig. 8.8. It consists of two amplifiers. The first amplifier accomplishes recharge sampling and imparts the majority of the gain to the signal. It is important to note that as much gain as possible must go into the first stage since the signal competes with this amplifier's own noise. Any subsequent amplification will have little effect on overall SNR. The second amplifier brings the signal to an acceptable level and adds DC offset.

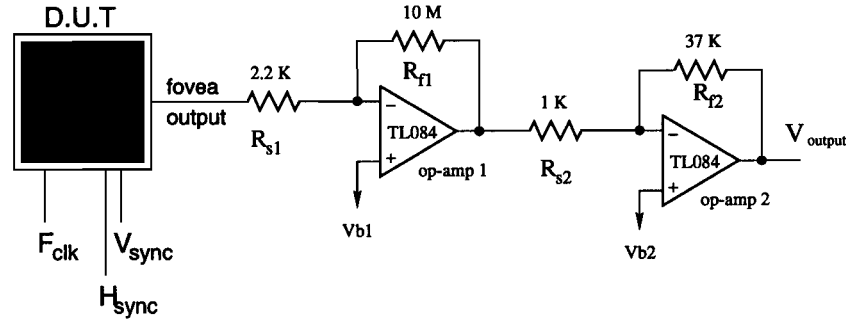


Figure 8.8: Off-chip circuitry for amplification of output from the fovea array.

The choice of particular amplifiers as well as the values of the resistors shown in Fig. 8.8, contributes to the performance of the fovea array. Op-amp 1 should have very low input noise, and be fast enough to reproduce the signal at acceptable frame rates. Larger values of resistor  $R_{f1}$  provide more gain to the signal charge, however they also slow the system down considerably. A value of  $10\text{ M}\Omega$  was found to yield good results. The values for the resistors in the second stage are chosen to provide adequate gain to the signal, and are not as critical as in the first stage.

Sample outputs from the fovea array using the readout circuitry described above, are shown in Fig. 8.9. Fig. 8.9a shows one complete frame of the output when a screw driver tip is imaged by the fovea array. Note that the dark reference bars described in Chapter 5 and Chapter 7 show up clearly in the waveform. These provide excellent ‘landmarks’ by which the general operation of the array may be judged on the oscilloscope. Fig. 8.9b is a closeup of one complete row of the waveform in Fig. 8.9a. Note the presence of spatial noise in the form of an alternating ‘high’, ‘low’ pattern every second column. The spatial noise performance of the fovea array will be examined in detail in Section 8.5.1.

As illustrated in Fig. 8.8, the fovea array is clocked using the  $F_{clk}$  input. This signal line drives the on-chip shift registers described in Chapter 7. The fovea array is composed of 40 columns of 52 pixels each, therefore, for a frame rate of  $f_{frame}$ , the

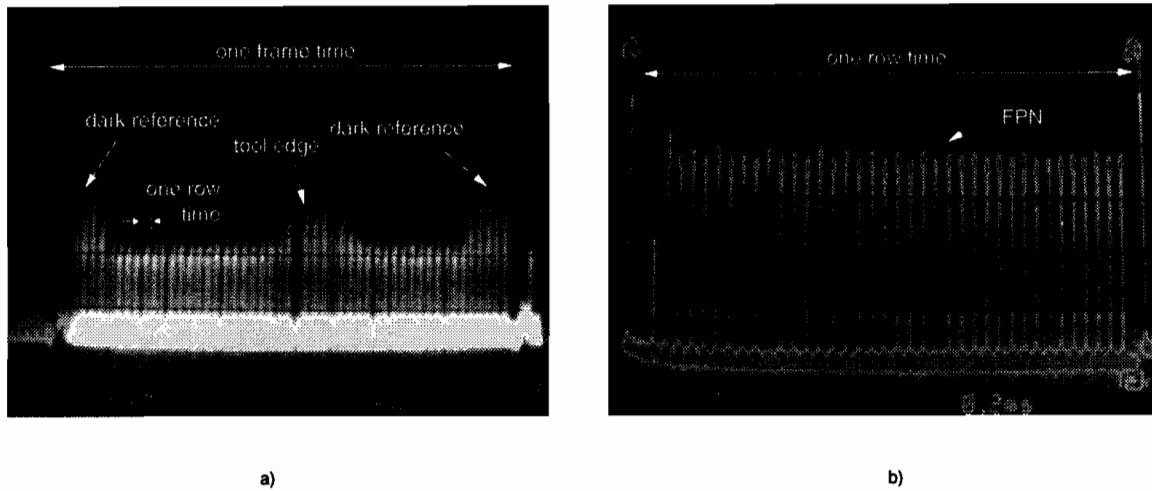


Figure 8.9: Oscilloscope traces of output waveforms for fovea array. a) One complete frame of pixels. b) One complete row of pixels.

$F_{clk}$  input must be clocked at a frequency,

$$f_{F_{clk}} = (40 + 1) \times (52 + 1) \times f_{frame} \quad (8.3)$$

Note that, as was the case with the periphery array, each time the X or Y selection registers are reset, a synchronization pulse is generated. This pulse accounts for the additional phantom row and phantom column in Eq. (8.3). As shown in Fig 8.8, these pulses are buffered off-chip ( $H_{sync}$ ,  $V_{sync}$ ) and are extremely useful both for monitoring array performance as well as for triggering an oscilloscope in order to lock the displayed waveform.

#### 8.3.4 Interface circuitry

In addition to the above described circuitry for driving the prototype imager, Analog to Digital (A/D) interfaces were developed to facilitate display of the image data. Two separate data gathering hosts were used in order to best observe the behavior of the sensor. For real-time effects, the sensor test circuit was interfaced to a Personal Computer (PC). Data used for more involved analysis of the sensor performance was captured using an HP e1430a A/D converter connected to a SUN SPARC station

over the VXI bus.

### **Interface to PC**

To interface the foveated sensor to a personal computer, a commercially available 8-bit semi-flash A/D converter was chosen. The outputs of this converter are read by the PC through the Centronix parallel printer port. The clocks for both the A/D and the foveated sensor are generated by the PC for proper synchronization. This setup provides for real-time operation at close to 30 frames/sec, however the PC-generated clock suffers from clock-jitter. This phase-noise affects both the integration time and the pixel sample point and therefore leads to increased uncertainty in the output data. Consequently, the system is useful for experimentation with real-time applications, but is too noisy to produce good static images.

### **Interface to SUN workstation**

In order to obtain high quality images, as well as to perform statistical analysis on the raw data, the sensor was interfaced to a SUN workstation using a 23-bit 10 MHz A/D converter card. With this configuration it is possible to sample over 500 consecutive frames of image data for subsequent processing and analysis. The output of the sensor readout circuitry is buffered by a discrete Bipolar Junction Transistor in order to drive the 50  $\Omega$  input impedance of the A/D card. The clocks for driving both the sensor circuitry and the A/D converter are generated using a stand-alone function generator producing clean waveforms free of phase-noise. This A/D card setup was used to gather all of the image data presented and analyzed in the remainder of this thesis.

## **8.4 Test of the periphery**

In this section, the test of the periphery array of the fabricated prototype is explained. Data gathered using the setup of Section 8.3.2 is analyzed; images of real scenes are presented; and the noise performance of the prototype system is discussed.

### 8.4.1 System calibration

The external circuitry described in Section 8.3.2 scales and formats analog data from the prototype sensor for input to a computer. In order to view the resulting image data however, it is necessary to fine tune the gain for each ring and adjust for offsets. A routine for accomplishing this task was written in the MATLAB programming environment [84], and involves a calibration procedure by which system offsets and gain mismatches are determined and corrected.

The first step in the calibration procedure involves nulling of system offsets. A block of image data is sampled in continuous time with the lens aperture closed. Since no light falls on the device while the data is being acquired, the resulting output dark image sequence  $d(r, \theta, t)$ <sup>7</sup>, serves as a reference containing all the offsets present in the system. Although temporal noise in periphery image data is usually low, it is important to have a very accurate dark image to obtain good results in subsequent processing. To remove temporal noise, the MATLAB routine averages  $T$  consecutive frames within the data block, producing the output frame,

$$\overline{d_t(r, \theta)} = \frac{1}{T} \sum_{t=1}^T d(r, \theta, t) \quad (8.4)$$

In addition to containing ring offsets,  $\overline{d_t(r, \theta)}$ , contains the dark pattern of the particular array and will be used in Section 8.4.4 to quantify fixed pattern noise (FPN) in the fabricated prototype.

Once a good dark image has been created, the next step in the calibration procedure is to determine scaling coefficients,  $\kappa(r)$ , for fine-tuning the gain of each ring in the array. As was discussed in Chapter 6 the transfer characteristic of photodiodes in the imager is relatively linear for small changes in exposure, and nonlinear for large changes. This effect leads to different scaling coefficients  $\kappa(r, e)$  at different exposures  $e$ . To determine  $\kappa(r, e)$ , it is necessary to obtain the response of the array when imaging a blank screen at  $E$  different exposures. Therefore, to calibrate the periphery imaging system, the opto-mechanical setup described in Section 8.3.1 was used to sample  $E$  data blocks  $K(r, \theta, t, e)$ , at intensities,  $P_e$ . The data block for  $e = 0$

<sup>7</sup>See Chapter 6 for a discussion of the notation used in this section.

corresponds to irradiance  $P_0 = 0$  and is therefore equivalent to  $d(r, \theta, t)$ . The last block is sampled at an irradiance halfway between  $P_0$  and  $P_{sat}$ , the saturation irradiance. The intensities for the remaining data blocks are distributed evenly between these two as follows,

$$P_e = \frac{P_{sat}}{2E}e, e = 1 \dots E \quad (8.5)$$

Each one of the  $E$  data blocks is processed by the calibration routine. To remove temporal noise,  $T$  subsequent frames in the data are averaged together as specified in Eq. (8.4), to produce the output frames,  $\overline{K_t(r, \theta, e)}$ , for the particular exposure. To remove variation in this image due to nonuniform illumination of the blank screen, the central pixels along each ring of the array are averaged together as follows,

$$\overline{K_{t,\theta}(r, e)} = \frac{1}{\theta_2 - \theta_1} \sum_{\theta=\theta_1}^{\theta_2} \overline{K_t(r, \theta, e)} \quad (8.6)$$

where  $\theta_1$  and  $\theta_2$  are chosen heuristically. The resulting intensity values  $\overline{K_{t,\theta}(r, e)}$  represent the overall response for a particular ring at the given exposure.

Once the  $\overline{K_{t,\theta}(r, e)}$  intensity values have been determined, ring gain values  $\kappa(r, e)$  can be calculated for a properly calibrated image. Gain values are normalized about the smallest ring in the periphery array as follows,

$$\kappa(r, e) = \frac{\overline{K_{t,\theta}(1, e)} - \overline{K_{t,\theta}(1, 0)}}{\overline{K_{t,\theta}(r, e)} - \overline{K_{t,\theta}(r, 0)}} \quad (8.7)$$

The final step in the calibration procedure is to average together all normalized gain coefficients for a given ring for the  $E$  exposures,

$$\overline{\kappa_e(r)} = \frac{1}{E} \sum_{e=1}^E \kappa(r, e) \quad (8.8)$$

This yields an approximation to the true nonlinear transfer characteristic which greatly reduces the complexity of the calibration process.

Once  $\overline{d_t(r, \theta)}$  and  $\overline{\kappa_e(r)}$  have been determined for a particular setup, fully calibrated image data vectors  $P(r, \theta)$  for the image sequence,  $I(r, \theta, t)$ , are produced according



to the following expression,

$$P(r, \theta) = \left( \overline{I_t(r, \theta)} - \overline{d_t(r, \theta)} \right)^T \Lambda_k \quad (8.9)$$

where,  $\Lambda_k$  is a diagonal matrix such that,

$$\begin{aligned} \Lambda_k(i, j) &= \overline{\kappa_e(r)}, i = j = r \\ &= 0, i \neq j \end{aligned} \quad (8.10)$$

Fig. 8.10 illustrates the calibration of actual image data taken from the fabricated prototype. Each horizontal line in the black and white images at the top of the figure represents data sampled from one ring in the periphery array. In the graphs at the bottom of the figure, these rings of data are plotted so that their relative scale and offsets may be compared. Fig 8.10a shows the original, uncalibrated data. Note that some rings have quite large amplitudes whereas others have very small amplitudes; in addition the relative offset varies quite a bit. In contrast, the plot of Fig 8.10b shows the image data after calibration, with  $E = 6$  and  $T = 400$ . Note that the amplitudes are all nearly equal, and the offsets have been eliminated. The precision of the calibration technique can be appreciated by comparing the region marked **A** in both plots.

Depending on the degree of noise which may be tolerated, the implementation of Eq. (8.9) can be simplified. Fig. 8.11 illustrates this point. All of the images have been normalized according to the procedure described above, but differ in the presence or absence of spatial and temporal noise. Fig. 8.11b and d were created by substituting  $I(r, \theta, 1)$  instead of  $\overline{I_t(r, \theta)}$  in Eq. (8.9) and therefore display temporal noise. Fig. 8.11c and d were created by substituting  $\overline{K_{t,\theta}(r, 0)}$  in place of  $\overline{d_t(r, \theta)}$  in Eq. (8.9) and therefore display FPN. Interestingly, these substitutions do not affect image quality significantly, and this is a reflection of the favorable noise performance of the periphery array.

It is important to note that the above-described calibration procedure, though somewhat involved, need only be performed once for a particular imager setup. Once the necessary offset and gain normalization coefficients have been determined the sensor can be used to image real scenes. Although not implemented in the present system, it is possible for a DSP engine (such as the TMS320C40) to perform the required

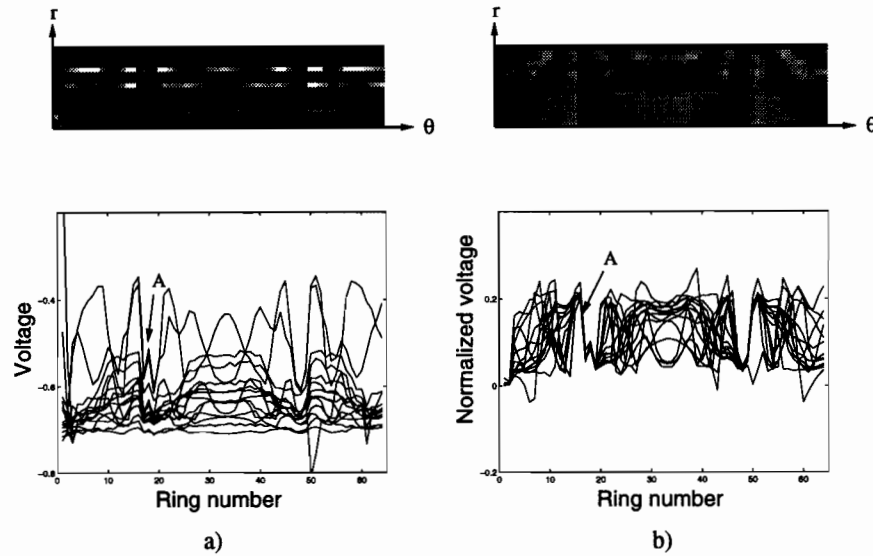


Figure 8.10: Illustration of image calibration. a) Uncalibrated image b) Calibrated image.

subtraction and multiplication operations of Eq. (8.9) in real-time, producing fully calibrated video images from the prototype sensor. Furthermore, the simplifications used to produce the images of Fig 8.11 can reduce the required processing considerably, thereby rendering the DSP solution even more attractive.

Finally, although it is possible to realize a precisely calibrated sensor setup using only off-chip analog electronics for scaling and offset correction, the best solution to this problem is to implement all of these functions directly on-chip. These functions may be performed directly at the pixel level even before data is read out, yielding a very accurate fully calibrated sensor.

#### 8.4.2 Example images

To illustrate the operation of the periphery part of the prototype sensor, image data was taken using the setup described in Section 8.3.1. The array was clocked at a rate of  $f_{master} = 41.5$  kHz which corresponds to a frame rate of 40 frames/sec. The data was input to the MATLAB routine described in Section 8.4.1 with parameters,  $T = 400$  and,  $E = 6$ ; the output data  $P(r, \theta)$  was then converted to graphical format for analysis.



Figure 8.11: Temporal and spatial noise in the periphery array. a) Noise free data b) Original data less FPN c) Original data less temporal noise d) Original data with noise.

Fig. 8.12 shows input images on the left and corresponding output images on the right. In Chapter 2, it was explained that the foveated mapping produces output images in log-polar coordinates, such that straight lines in the original scene are mapped to curved lines in the periphery output image. This effect is clearly visible in Fig 8.12a in which an array of parallel black bars in the original scene results in a group of curved bars in the output. The effect is also visible in Fig 8.12b where the same original image used in the software simulation of Fig. 2.7 is imaged by the fabricated prototype. Finally, Fig 8.12c shows how a wristwatch is warped by the foveated mapping. It is interesting to note that, although the global topology of the original scene is altered by the mapping, local topology does not change very much. In fact, the watch face in Fig. 8.12c is easy to recognize, even though the wristband has been transformed to a curved shape. The preservation of local topology is important because it means that interest features in the original scene may be recognized in the periphery output image and used to direct the gaze of an autonomous robot vision system [85].

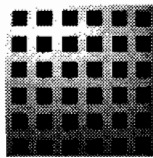
### 8.4.3 Imperfections in the output image

In the output image of Fig. 8.12c, there are some imperfections due to physical effects at the sensor level. The first of these imperfections are black lines in the image. These are visible along the left edge of the image, along the top of the image, and in the

top right corner. These lines indicate that no image information was sampled from corresponding photodiodes on the device. In fact, as explained in Appendix B, the photodiodes in question are not accessible in the fabricated prototype. Therefore the lines will show up in any image scanned from the sensor, and are clearly visible in Fig. 8.12a and b.



a)



b)



c)

Figure 8.12: Example images from the periphery array. a) 12.5 mm parallel bars b) 12.5 mm blocks c) Wristwatch. Original scene is shown on the left, output of the array is shown on the right.

The second imperfection which is visible in the output image of Fig. 8.12c, is a row of pixels slightly brighter than all others. This same line is slightly visible in Fig 8.12b, and less so in Fig 8.12a. The line suggests a possible error in the calibra-

Table 8.2: Factors limiting the speed of the periphery array.

Factor	$f_{P_{clk}}$ (kHz)	Frame rate (frames/sec)
$P_{sync}$ breakdown	480	7385
op-amp transient	250	3846
off-chip circuitry	13.4	206

tion procedure discussed in Section 8.4.1, however the source of the defect is in fact hardware related. The origin of this anomaly becomes apparent by observing that the order in which the images were taken is the order in which they are shown: i.e. the effect becomes worse over time. As was explained in Section 8.3.2, data is read from the periphery array by sampling voltage transients from transimpedance amplifiers. Any subtle change in these transients or the sampling instant will cause a change in the offset and scaling factor of the particular ring, requiring complete recalibration of the sensor setup. Lack of proper calibration results in lines in the output image with varying brightness and contrast. This effect can be alleviated using more complex off-chip circuitry, however, the best solution is the integration of all sampling circuitry directly on the same substrate as the sensor matrix.

#### 8.4.4 Performance measurements

##### Speed of operation

The periphery array benefits significantly from a parallel output architecture, and therefore displays very high speeds of operation. To quantify this performance, it is useful to examine limitations in the prototype system. These are tabulated in Table 8.2 and discussed below.

The on-chip shift registers are the fastest components in the imager system. In the prototype imager,  $f_{P_{clk}}$  was increased to 480 kHz before the output synchronization pulse  $P_{sync}$  began showing signs of decay. This corresponds to a maximum frame rate of 7385 frames/sec<sup>8</sup>, and this represents the maximum speed of operation of the

<sup>8</sup>It is believed that the on-chip circuitry operates above this rate and that the limiting factor is the digital output pad.

device.

In the present system, however, off-chip amplifiers reduce this speed. The minimum time between consecutive pixels at the output of the 15 transimpedance amplifiers was measured to be  $4\ \mu\text{s}$ , corresponding to a maximum frame rate of 3846 frames/sec. The final limiting factor in the present setup however is the sampling and multiplexing system of Fig. 8.4a. The output of the multiplexer was found to decay when the system was clocked at 214 kHz<sup>9</sup>, corresponding to a maximum frame rate of 206 frames/sec.

Since the limitation in the present system appears to be off-chip circuitry, it is conceivable that future versions of the sensor incorporating on-chip amplifiers and an on-chip multiplexer might be made to operate well above 1000 frames/sec. It is important to remember however, that the sensitivity of an integrating sensor is inversely related to its frame rate. Therefore, high frame rate operation would require either very low noise signal processing or compensating artificial illumination of the perceived scene.

### Quantification of temporal and spatial noise

To investigate the noise performance of the periphery array, a MATLAB routine implementing the methods discussed in Chapter 6 was used to calculate the signal-to-fixed-pattern-noise-ratio, (SFPNR) and the dynamic range with respect to temporal noise (DR), for each ring in the fabricated prototype. These values were calculated using the same 400 consecutive frames of raw data used to generate the images of Section 8.4.2. The results of this procedure are graphed in Fig. 8.13. It is interesting to note that both spatial and temporal noise performance improves with ring number. This phenomenon is also evident in Fig. 8.11 and is probably due to the fact that the saturation output is greater for rings which contain larger photodiodes.

Although trends in the data are evident, some anomalies do exist<sup>10</sup>. These are

<sup>9</sup>Clocking the overall system at 214 kHz corresponds to clocking the  $f_{CLK}$  line at 13.4 kHz since  $f_{CLK} = f_{master}/16$ .

<sup>10</sup>The most intriguing anomaly occurs for ring number 8: neither the data points for DR or SFPNR appear to fit the general trends of Fig 8.13. Although the exact source of this difference has not been investigated, it appears likely that the value of the resistance ( $120\ \Omega$ ) was not recorded correctly.

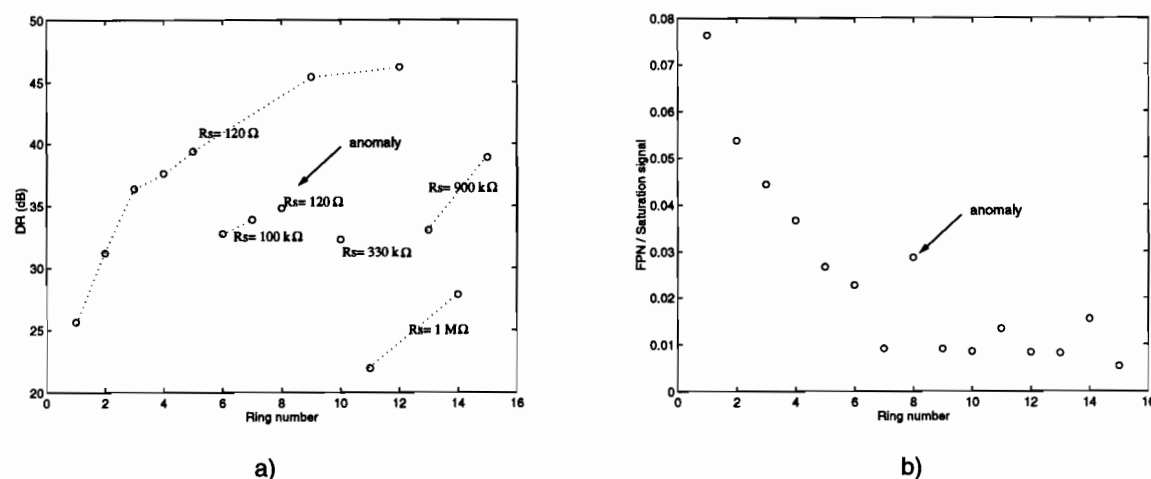


Figure 8.13: Noise in the prototype system. a) Graph of dynamic range of sampled image data with respect to ring number showing effects of temporal noise. b) Graph of SFPNR output of sampled image data with respect to ring number showing effects of spatial noise.

especially evident in Fig. 8.13a and appear to be related to the value of the resistance  $R_s$  (see Fig. 8.4). Even though the trend of increasing DR with ring number is preserved, lower values of  $R_s$  appear to give better noise performance. Although a detailed noise analysis has not been performed, the discussion of Chapter 6 suggests that the presence of this noisy resistance in the signal path has a significant effect. Future versions of the sensor would benefit significantly from a complete theoretical noise analysis.

### Power consumption

The fabricated prototype dissipates dynamic power through on-chip digital scanning circuitry as well as in its input and output buffers. This power can be determined by measuring the supply current of the device. To accomplish this measurement, a  $10\ \Omega$  resistor was placed between the power supply and the sensor. The voltage waveform across this resistor was amplified using an instrumentation amplifier and then sampled using an A/D converter. The integrated power was then calculated using MATLAB. The periphery array was clocked at 2.6 kHz corresponding to a frame rate of 40 frames/sec. Power was measured at two different supply voltages and is tabulated in Table 8.3.

Table 8.3: Power consumption of periphery array.

Supply (V)	Power ( $\mu$ W)
5	410
3.3	35

Note that these numbers are for on-chip digital electronics only. Power dissipation in off-chip sampling and amplification circuitry is considerably higher due to the use of discrete amplifiers. It is believed that off-chip analog processing could be incorporated on-chip using a switched-capacitor approach which could lead to savings in power dissipation for these functions.

## 8.5 Test of the fovea

In this section, the test of the fovea array of the fabricated prototype is explained. Image data sampled using the setup described in Section 8.3 is presented and analyzed. The performance of the fovea array is then quantified by measuring its speed of operation, power consumption and electrical noise.

### 8.5.1 Analysis and processing of the raw image data

As was explained in Chapter 6, a number of noise sources conspire to reduce the quality of image data in real arrays; in addition limitations in the output amplifier discussed in Chapter 5, contribute further degradation in overall image quality. To investigate these effects, data was sampled from the fovea array while it was clocked at 65.2 kHz corresponding to a frame rate of 30 frames/sec. This data was then converted to an output image and the results are shown in Fig. 8.14a. For illustrative purposes, voltage values for one row in the image (corresponding to the cross-sectional line A-A'), are plotted in Fig 8.14b. The various imperfections in fovea output are discussed below.



## Chapter 9

# Conclusion

### 9.1 Towards efficient vision sensors for autonomous robots

In the future, autonomous mobile robots with some degree of artificial intelligence will be expected to operate unsupervised in hostile environments. The success of these machines hinges on the efficient use of limited available on-board resources such as computers, sensors and power sources. A conservative design strategy calls for the optimization of all components of the system such that the sum of the parts is compact and energy-efficient. When applied to future vision sensors, this methodology calls for highly integrated, low-power devices with reduced data overhead due to a careful focus of sensor attention.

In this thesis, a first step towards the realization of such vision sensors was undertaken. A CMOS image sensor implementing a biologically motivated space-variant sampling scheme known as the foveated mapping was designed and tested. In this chapter, the work is summarized and discussed.

### 9.2 Review of the present system

The prototype CMOS foveated image sensor was implemented in an industry standard  $1.2\ \mu\text{m}$  mixed-signal CMOS process<sup>1</sup>. The device is mounted in a 44 pin Leadless Chip Carrier (LCC), and measures  $4.8\ \text{mm} \times 4.8\ \text{mm}$  on the die. The device contains two separate image sensing arrays. The fovea array consists of  $52\ \text{rows} \times 40\ \text{columns}$

---

<sup>1</sup>Nortel's 2 metal, 2 poly CMOS4S process [26].

of PPS pixels with a  $9.6\mu\text{m}$  pitch. The periphery array consists of  $64 \text{ rays} \times 16$  columns of PPS pixels. Pixels in the periphery array are distributed on a log-polar grid and grow in size with eccentricity from the center. Scanning of both of these arrays is accomplished using dynamic logic shift registers which have been integrated directly on the same substrate as the respective sensor matrices.

The imager is supported by external circuitry for signal readout and conditioning. Readout of the fovea array is accomplished using a high gain transimpedance amplifier followed by a voltage gain stage. Readout of each ring in the periphery is accomplished using a high gain transimpedance amplifier followed by a sample and hold circuit. All 16 rings are read simultaneously, and these are then multiplexed using external analog multiplexers. All off-chip components use commercially available integrated circuits.

Image data is read from the above described system into a digital computer using two possible configurations. The use of a PC and 8-bit A/D converter allows image data to be displayed in real time on the video monitor. This is a reflection of the considerable reduction in image data afforded by the foveated mapping. The use of a SUN workstation and 23-bit A/D converter card on the VXI bus allows large blocks of image data to be gathered in real-time for later processing. This facilitates digital signal processing of the image as well statistical analysis of the system noise performance.

The present system functions well at 30 frames/sec. The periphery array yields good quality images with high dynamic range. The fovea array produces images with lower dynamic range but which are still useful for imaging of high-contrast objects such as tools. The sum of all pixels in the fovea and periphery arrays amounts to 3104 data points per frame. If the entire sensor area were covered in pixels the size of those in the fovea, the resultant uniform resolution sensor would output 250 000 pixels per frame. Therefore a reduction in image data of roughly 80:1 has been realized. The use of a standard CMOS process means power consumption is likewise reduced. Both the fovea and periphery array driving circuitry consume less than 1 mW of power at 30 frames/sec operation. Frame rates in excess of 1000 frames/sec are supported by the scanning circuitry of both arrays, but are limited by the use of slow output amplifiers. Modifications to the present system can be made for increased performance, and these

will be described in Section 9.4.

### 9.3 Discussion of key issues

A number of key points surfaced during the discussions in the previous chapters. A review of these is presented here to provide an overview of the ideas presented.

In Chapter 2, the foveated mapping was examined. It was shown to be motivated by a study of biological evidence concerning the structure of the human visual pathway. It was further found to possess useful properties such as the realization of a high degree of image data reduction as well as rotation and scale invariance. The hybrid model was developed in order to allow for implementation of the foveated mapping as an integrated circuit. Although this model results in a loss of 1/5 of the original image information, it makes possible very efficient realization of the periphery array in standard CMOS.

Chapter 3 and Chapter 4 provided background for the study of CMOS imagers. Various silicon phototransducers were examined in Chapter 3, and the photon flux integration mode of operation was introduced as an efficient method for signal extraction with photodiodes. Phototransducers were next extrapolated to image sensors in Chapter 4. A comparison of CCD and CMOS imager technology was presented which showed that CCD's consumed far more power and were limited due to the need for very high charge transfer efficiency (CTE). Finally, a survey of the three types of CMOS imager: smart pixel, active pixel, and passive pixel was presented.

The CMOS PPS imager was examined in detail in Chapter 5. It was shown to be composed of a sensor matrix, scanning circuitry and an output amplifier. Progressive-scan of the sensor matrix can be accomplished in an efficient manner using dynamic logic shift registers. Care must be taken to scan the array properly such that charge-sharing effects do not corrupt the signals in high resolution PPS arrays. Through simulations performed on an  $8 \times 8$  test circuit, the limitation on maximum frame-rate was shown to be the off-chip output amplifier of the device. Finally, the use of extra pixels physically covered in metal known as dark reference pixels was instituted to facilitate test of the fabricated prototype.

Chapter 6 expanded on the discussion of Chapter 5 by examining some nonideal effects in PPS imagers. The saturation output  $V_{sat}$  was shown to be limited by the threshold voltage of NMOS transistors and by clock feed-through. It was explained that two noise sources plague solid-state imagers: temporal noise arising due to thermal fluctuations in on-chip resistances, and spatial noise (FPN) mostly due to variation in process parameters. The best way to eliminate these is through differential readout methods, however subsequent DSP processing is especially useful for removing FPN. Two useful metrics are the dynamic range with respect to temporal noise, and the signal-to-fixed-pattern-noise-ratio, (SFPNR). These were introduced and later used in Chapter 8 to quantify the noise performance of the prototype sensor. The chapter concluded with a discussion of charge leakage effects which are manifested in PPS imagers as blooming and smear.

A discussion of the design of the CMOS foveated sensor prototype was presented in Chapter 7. The minimum pixel size of PPS type imagers makes this technology the ideal choice for the fovea array. For optimum resolution substrate contacts in the form of a  $p^+$  diffusion ring were placed outside of the array. Although this design compromise lowers the resistance to blooming, it is necessary to satisfy the requirement of maximum spatial resolution for the fovea. It was shown that implementation of the periphery array was best achieved through the use of specially developed layout software. In addition, the use of round photodiodes and a modular architecture greatly simplifies the layout on a standard CMOS design grid. It was decided that amplifiers for both arrays would be off-chip in order to achieve maximum observability of the integrated circuitry. This would facilitate testing of the fabricated prototype and allow information to be gathered for future implementations of the sensor.

Chapter 8 presented an in-depth discussion of the techniques used to evaluate the fabricated prototype along with results from actual experiments performed in the lab. Test apparatus was developed in order to properly characterize the device. This included opto-mechanics to allow the focusing of real-world images on the die surface, as well as support circuitry for sampling data from the device, and A/D interfaces for gathering of image data. Test of the periphery and fovea arrays was facilitated through the use of frame and line synchronization pulses, and dark reference pixels.

It is well known that the first gain stage in a series of such stages should be the one with the most gain for good noise performance. This rule of thumb was used in the design of the fovea output circuitry in order to reduce noise in the output of the fovea array. A calibration routine was developed to fine tune gain adjustments on the output of each ring of the periphery array. In addition, FPN in the fovea array was removed by subtracting a dark frame from all images. Both of these procedures were performed using digital signal processing (DSP), and could conceivably be implemented in real-time using special purpose DSP hardware.

## 9.4 Suggestions for future work

In addition to the issues reviewed in the last section, a number of suggestions for possible future work were noted, and these will now be discussed.

In the immediate future, integration of the prototype imager into a mobile robot system is the next step towards understanding the potentials of this technology. The performance outlined above appears useful for general purpose applications, however tests in the laboratory on a mobile robot should reveal important information. For example, it would be useful to know whether the noise performance reported for the fovea and the periphery arrays is adequate for typical mobile robot tasks. In addition, experiments performed with a robot in real-time could be used to determine ideal maximum frame-rates for future designs. Finally, only through integration into a real system can the savings in power and system mass of this approach be appreciated and properly evaluated.

Significant improvements in image quality and sensor performance could be realized through a detailed redesign of the present system. Bringing all of the off-chip analog support circuitry onto the sensor substrate is the next logical step in realizing these improvements. The use of on-chip charge amplifiers would make low-noise reading of the fovea array possible, and allow for proper scaling of the periphery array outputs before they were brought off-chip. Careful redesign with attention to such details as minimizing clock feed-through and parasitic capacitances could result in much higher SFPNR. In addition, a complete theoretical noise analysis of the signal

paths in the imager could yield insight into design modifications which would improve the dynamic range of the fovea and periphery arrays. Finally, a detailed analysis of the interaction of analog and digital circuitry on the same substrate and the use of differential readout techniques could lead to the elimination of various anomalies reported in the test results for the fovea and periphery arrays.

Once the present system has been evaluated in practise and improvements in basic image quality have been made, the next step towards better imagers is a detailed examination of changes in the architecture of the sensor. Improvements along these lines include expanding the size of the fovea from  $52 \times 40$  to as much  $100 \times 100$ , expanding the resolution of the periphery array, as well as the elimination of blind spots in the fovea. The use of more advanced CMOS processes with smaller feature sizes could significantly increase the resolution in the fovea and periphery, while still keeping the overall sensor area small. The use of a standard CMOS process makes the integration of on-chip A/D converters quite feasible [75], [86], and would facilitate interface of the sensor to external digital systems. Some degree of intelligence could be incorporated into the sensor in the form of circuitry for monitoring and adjusting exposure, frame-rate and power consumption. In addition, an investigation into the possibility of integrating further analog and digital processing directly on the sensor substrate could be undertaken. Higher level machine vision algorithms could be implemented much more efficiently with the increased parallelism available to such an integrated processing system. Finally, present research indicates that color CMOS imagers will soon become available. Future prototypes with color fovea arrays could prove quite useful in autonomous mobile robot tasks such as the manipulation of color-coded tools and navigation in environments containing distinctly colored features.

## Appendix A

### Optical test setup

The image sensors which are used to take pictures of our macroscopic world are in fact microscopic devices. For example, every pixel in the fovea array of the prototype CMOS foveated sensor is only  $9.6\text{ }\mu\text{m}$  wide; the diodes on the test chip discussed in Section 8.2 are only  $40\text{ }\mu\text{m}$  wide. Although the camera lens system of Section 8.3 could in principle be used to test the photoelectric response of a group of such minute structures, localizing this response for each individual transducer is extremely difficult. What is needed is a type of optical *probe* – analogous to the oscilloscope probes used in electronics – which would permit the selected illumination of tiny areas of the die surface.

Fig A.1a illustrates the topology of an optical probing system such as the one described above. A system of cemented achromat lenses are used to image incoherent light of wavelength  $880\text{ nm}$  from an Infrared Light Emitting Diode (IR LED). The diffraction-limited radius  $\Delta l$  of the resulting spot on the surface of the test device (D.U.T.) is given to a first order approximation, by the following expression [83],

$$\Delta l = 1.22 \frac{f\lambda}{D} \quad (\text{A.1})$$

where  $\lambda$  is the wavelength of the incident radiation,  $f$  is the focal length of lens No. 1, and  $D$  is the effective diameter of lens No. 1. Nominal values for these parameters in the experimental setup were,  $f = 10\text{ mm}$ ,  $D = 5\text{ mm}$ , and  $\lambda = 880\text{ nm}$ . This implies that the minimum spot size which can be obtained by the system is roughly  $4.3\text{ }\mu\text{m}$  in diameter. In practise, the observed spot size was larger due to aberrations at lens No. 1, as well as accumulated aberrations in the imaging optics discussed below.



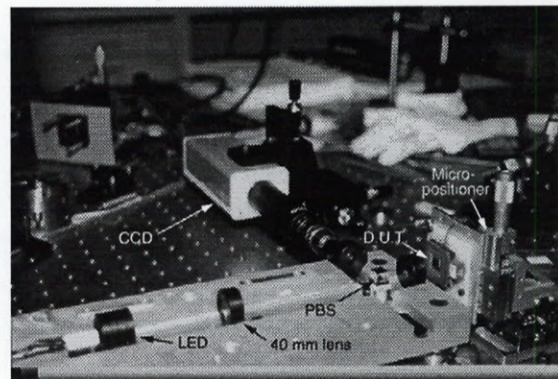
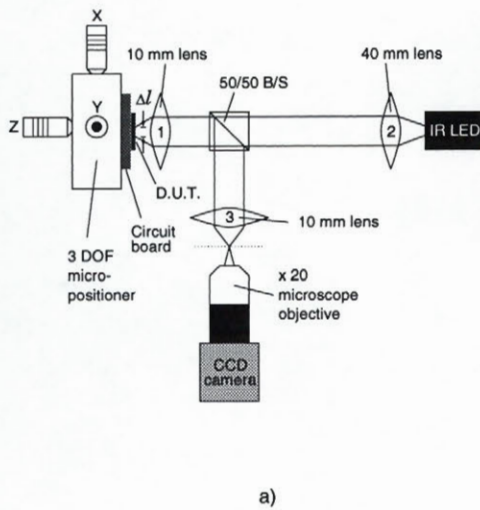


Figure A.1: Experimental optical setup. a) Schematic diagram. b) Photograph of experimental setup.

The test device was mounted on a three Degree-Of-Freedom (DOF) micro-positioner, allowing for submicron accurate  $X$  and  $Y$  translation in the image plane, and  $Z$  translation for fine tuning of the spot size. As illustrated in Fig. A.1a, a 50/50 Beam Splitter (50/50 B/S) was placed in the path of the main optical train to allow imaging of the die surface. A  $20\times$  microscope objective provided reasonable magnification of the image such that it could be viewed using a standard CCD camera.

Fig A.1b shows a photograph of the experimental setup in the lab. An optical base-plate developed by the McGill Photonic Systems Group[87] and based in part on research conducted at AT&T[88] was used to simplify positioning of components in the system. Using the base-plate, it was possible to fine tune the size of the probe spot by changing the position of lens No. 2 along the optical path. As can be seen in the figure, the test device was mounted on a circuit board which is then secured to the micro-positioner. All of the components in the system were secured to the optical table using 1/4-20 (6.25 mm) screws for maximum mechanical stability.

With the above-described system it was possible to focus a localized spot of minimum width approximately  $20\text{ }\mu\text{m}$  directly on the surface of the test device. This



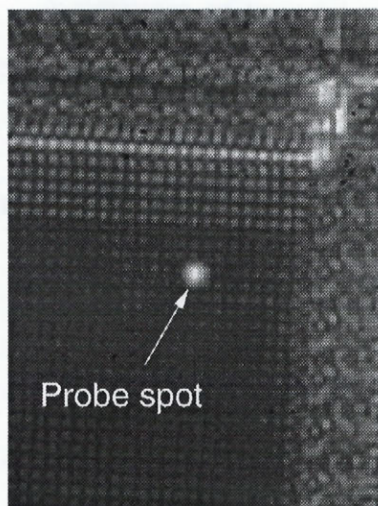


Figure A.2: Representative image of fovea array with light probe spot projected on it.

spot could then be translated in the plane of the device with sub-micron accuracy. In addition, the exact location of the spot was readily determined since the illuminated surface of the die was displayed on a television monitor. Fig. A.2 shows a typical frame-grabbed image as might be produced by the experimental setup of Fig. A.1. In this image, the individual pixels of the fovea array are visible. Note that the spot covers an area roughly equal to 2 pixels in diameter. During the experiment, the spot was swept along a row or column of the array by translating the device relative to the optics using the micro-positioner. The output of the array was observed on the oscilloscope and compared with the known location to determine whether the array was functioning correctly. As will be described in Section B.1, this system proved invaluable in determining and correcting defects in the fabricated prototype of the CMOS foveated sensor.

## Appendix B

# Microsurgery

When the original fabricated prototype of the CMOS foveated sensor was received from the foundry, rudimentary tests were run to verify proper functioning of the device. At that time serious problems were discovered with both the periphery and fovea arrays even though both of these exhibited excellent performance in simulation. Rather than starting from scratch with a completely new design, it was felt that much could be learned through a systematic attempt at rooting out the source of these problems. As will be explained below, extensive investigation not only led to the cause of the malfunctioning, but made possible the salvaging of the prototype itself. In fact, the results presented in Chapter 8 originate entirely from first-run prototypes which have had defects corrected through microsurgical techniques. These defects, and the procedures used to correct them are described separately for the fovea and periphery arrays in the sections which follow.

### B.1 Correction of periphery array defects

Tests performed on the fabricated prototype showed an error in the functioning of the periphery array. When the on-chip shift registers for scanning the periphery array were clocked using the  $P_{clk}$  input pad, the expected end of frame pulse at the  $P_{sync}$  output pad was not observed; instead the  $P_{sync}$  line was stuck at a high output level (5 V). This meant either that the output pad was malfunctioning, or worse, that the shift registers themselves were not responding. Optical point probing experiments using the apparatus described in Appendix A showed that all selection MOSFETs in



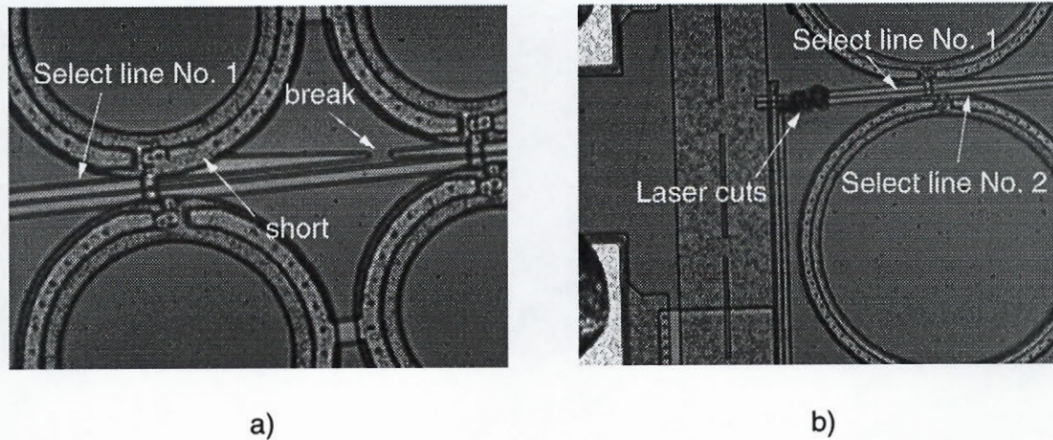


Figure B.1: Photomicrographs illustrating microsurgery on periphery. a) Close-up of vicinity of deformed select line. b) Close-up of vicinity of laser cuts on select lines No. 1 and 2.

the periphery array appeared to be turned off. This suggested that all of the on-chip shift register sections had somehow been reset to zero, indicating a possible short to  $V_{ss}$  at some point in the digital circuitry.

Such a defect might have occurred due to an error in the layout process. Were this the case, however, it would have manifested itself during simulation of the extracted layout. Since the extracted layout had performed as expected, an error in the initial design was ruled out. Nevertheless, it was still possible that an error during processing at the foundry might have caused a similar defect. The fabricated prototype was placed under the microscope and a detailed examination of the layout was begun. The results of this investigation were quite surprising.

Fig. B.1a, shows a photomicrograph of the defect which was uncovered through visual inspection. As can be seen in the figure, a portion of the digital select line for ray No. 1 is twisted out of shape near ring No. 10. This deformation creates a physical break in the line such that all photodiodes on ray No. 1 which are also on rings inside ring No. 10 are no longer accessible. More importantly, however, the twisted shape overlaps with the  $p+$  guard ring which surrounds the photodiode on ring No. 10 and ray No. 1. Since both of the lines are in the same metal layer, and the guard ring is kept at the substrate potential, a short between the digital select line and  $V_{ss}$  results.

It was felt that this defect might be the source of the problem. As was explained in Chapter 5, the on-chip selection circuitry consists of self-resetting shift registers. In normal operation, when all of the select lines go low, a distributed NOR gate generates a high reset signal at the input to the shift register. On the next clock pulse, this new high bit causes the NOR gate to drive the reset line low; consequently reset of the registers should create a distinctive low-high-low output pulse which can be observed at the  $P_{sync}$  output pad of the array. As was mentioned above, however, the  $P_{sync}$  output of the fabricated prototype was stuck high. Tests at very slow clock frequencies showed that the  $P_{sync}$  line started low right after power up and then became stuck at high after less than 65 clock cycles. It was postulated that the short on select line No. 1 was driving the output of the corresponding stage of the shift register low. Therefore, the select bit generated by the NOR gate would never be passed on to subsequent stages in the shift register since the output of the first stage was stuck low. Since all the select lines would remain low, the output of the NOR gate would always remain high. This would explain why the  $P_{sync}$  output pad started off low and then went high: on power-up, some of the dynamic logic shift registers could contain ones, and some of them zeros; within 65 clock cycles however, these would have been shifted out and replaced by zeros due to the bad first stage, thereby causing the output of the NOR gate to go high as was observed.

To verify this hypothesis and possibly correct the source of the problem, samples of the fabricated prototype were shipped to the laboratory of Prof. Glenn Chapman at Simon Fraser University in British Columbia. Researchers there were able to use a high power laser to perform microsurgical cuts on the device. As illustrated in Fig. B.1b, laser cuts were performed on digital select lines in order to isolate the shift register circuitry from the shorted metal layers. Even though select line No. 2 was not shorted to ground, the proximity of these lines ( $2\ \mu\text{m}$ ) made cutting them separately too difficult.

Once the cuts had been performed, the researchers tested the operation of the shift registers in their lab. The results were extremely encouraging. In most of the post-operative samples, the  $P_{sync}$  output now appeared to behave correctly, producing a synchronization pulse after 64 clock cycles on the  $P_{clk}$  line. The samples were sent

back to McGill, where the array was then tested to see whether image data could now be read from the sensor matrix; all of the post-operative devices performed well, thereby confirming the original hypotheses.

As was observed in Chapter 8, the periphery array now produces useful images. Due to microsurgery, however, select lines No. 1 and No. 2 are lost, and all corresponding pixels are no longer accessible. This results in black lines in the output image for rays No. 1 and No. 2.

In order that the error described above be avoided in the future, an investigation into its origin was launched with the help of officials at the Canadian Microelectronics Corporation. In the end, the problem was attributed to computer error: a software tool used during translation of the CIF file for the entire design was found to truncate database objects which contained more than a set number of vertices. This caused the elimination of a vertex of the affected digital select line which resulted in the deformation shown in Fig. B.1a. The problem has since been corrected, and is not expected to affect future designs.

Further inspection under the microscope revealed that the contact between the analog bus for ring No. 16 and the line to the corresponding analog output pad was absent. This defect means that all photodiodes on ring No. 16 are not accessible, and shows up as a black line in the output image for ring No. 16. This defect is unrelated to the ones discussed above, and can be avoided in future designs.

## **B.2 Correction of fovea array defects**

As was discussed in Chapter 5, correct scanning order is crucial to the operation of high resolution PPS arrays. In order that charge-sharing effects not corrupt the minute signal charge of the pixels in the fovea array, column scanning must be performed at the pixel clock rate, and row selection at  $1/N$  times this rate, where  $N$  is the number of columns in the sensor matrix. Due to a design error, scanning order in the fovea array of the fabricated prototype is reversed. Therefore, in order to obtain useful images from the fovea of the sensor, it was necessary to correct this defect

using microsurgery.

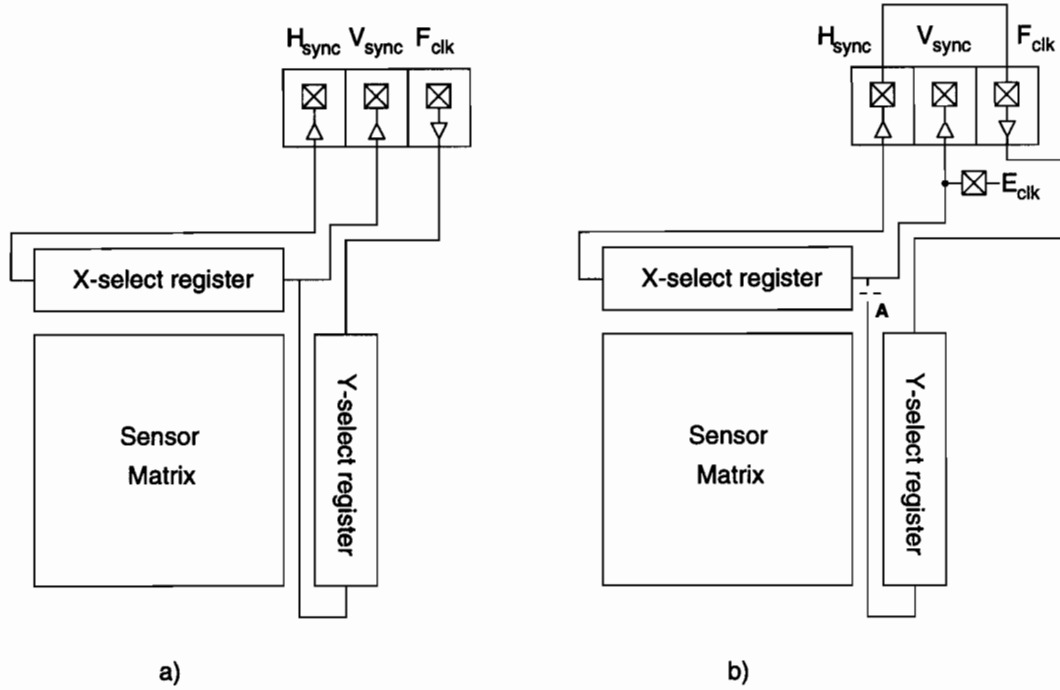


Figure B.2: Schematic illustrating surgical procedure for fovea array. a) Diagram of unaltered device. b) Diagram of altered device.

Fig. B.2 illustrates the procedure which was used to correct the defect. In the unaltered prototypes (Fig. B.2a), the buffered clock input  $F_{clk}$  drives the y-select register directly. The reset pulse of the y-select register drives both the x-select register and the output pad for the vertical synchronization pulse signal,  $V_{sync}$ . The reset pulse from the x-select register drives the output pad for the horizontal synchronization pulse signal,  $H_{sync}$ . Unfortunately, this arrangement means that, in the unaltered device, rows are scanned quickly whereas columns are scanned slowly. For correct operation, this order should be reversed.

As illustrated in Fig. B.2b, it is possible to accomplish this reversal using the building blocks already available on the die. Cutting the output line from the y-select register at the point marked **A** will leave the  $V_{sync}$  line floating. If this line can



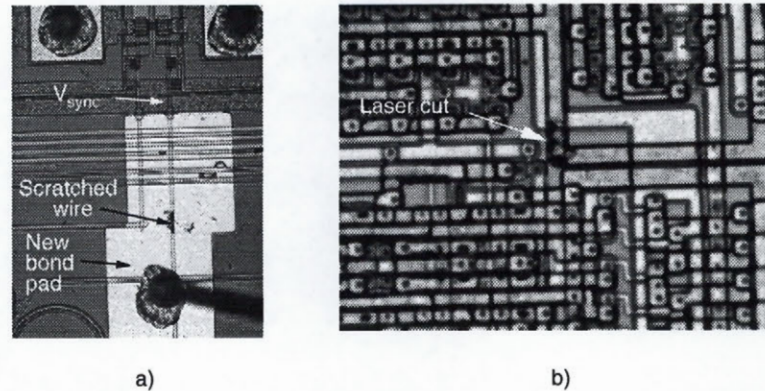


Figure B.3: Photomicrographs illustrating microsurgery on the fovea array. a) A new bond pad was patterned in aluminum on the die surface. b) The  $V_{sync}$  line was then cut using a high power laser.

be driven by an external clock,  $E_{clk}$ , it could be used to clock the x-select register directly. The reset pulse from the x-select register,  $H_{sync}$  could then be connected to the  $F_{clk}$  input. In this way, the scanning order of the array is reversed as required, with the  $E_{clk}$  line as the main clock line<sup>1</sup>.

The microsurgery to accomplish the rewiring described above was more complicated than that used to fix the periphery as described in Section B.1. Before the necessary cut at location A in Fig. B.2b could be made, an unbonded die would be specially prepared to allow access to the floating  $V_{sync}$  signal line. Fig. B.3a illustrates this procedure which was performed with the help of Prof. I Shih and Philips Laou in the McGill Device Physics Laboratory in the Department of Electrical Engineering. To obtain access to the  $V_{sync}$  line, the corresponding metal trace was exposed by scratching away the surface passivation glass with the tip of a die probe. As can be seen in the figure, a rectangular section of aluminum was then patterned directly on the surface of the die in order to form a new bonding pad. Since the pad was formed directly above the exposed metal line, a contact to the  $V_{sync}$  signal line was created. Four dice were prepared in this fashion and sent to Prof. Chapmann's Laboratory at Simon Fraser University to complete the procedure.

<sup>1</sup>It is important to note that the  $E_{clk}$  line in the altered device will not be buffered or protected from Electro-Static Discharge (ESD) since it is not driven by a digital input pad. For this reason great care must be taken in handling the altered devices.

Fig. B.3b illustrates the second part of the operation in which the laser cut corresponding to location **A** in Fig. B.2b is visible. This cut was more difficult to make than the one described in Section B.1 however 4 successful devices were realized. Once the cuts were made, the devices were returned to McGill where they were bonded in 44-pin Leadless Chip Carrier (LCC) packages for testing. As is evident from the results presented in Chapter 8, the fovea arrays of these prototype sensors perform well.

The success of the procedures described above illustrates the importance of exhaustive experimental testing of VLSI designs. Although, in principle, digital circuitry should exhibit similar performance to that which was observed in simulation, defects in device processing may still arise in first-run prototypes. Rather than being scrapped, the affected devices should be scrutinized using all available techniques because they hold valuable information for the success of future work.



## Bibliography

- [1] A. Vladimirescu, K. Zhang, A.R. Newton, D.O. Pederson, A. Sangiovanni-Vincentelli. *SPICE User's Guide*. University of California, Berkeley, 1983.
- [2] R. Wodnicki. A foveated image sensor in standard CMOS technology. In *Proceedings of the IEEE 1995 Custom Integrated Circuits Conference*, pages 15.5.1-15.5.4.
- [3] J. Van der Spiegel, G. Kreider, C. Claeys, I. Debusschere, G. Sandini, P. Dario, F. Fantini, P. Bellutti, and G. Soncini. A foveated retina-like sensor using CCD technology. In C. Mead and M. Ismail, editors, *Analog VLSI and Neural Network Implementations*, DeKluwer Publishers, Boston, MA, 1989.
- [4] M. Bolduc, and M.D. Levine. A foveated retina for robotic vision. In C. Archibald, and P. Kwok, editors, *Research in Computer and Robot Vision*, World Scientific, River Edge, NJ, 1985.
- [5] P.M. Daniel, D. Whitteridge. The representation of the visual field on the cerebral cortex in monkeys. *Journal of Physiology*, 159:203-221, 1961.
- [6] S.A. Talbot, W.H. Marshall. Physiological studies on neural mechanisms of visual localization and discrimination. *American Journal of Ophthalmology*, 24:1255-1263, 1941.
- [7] D.H. Hubel, T.N. Wiesel. Uniformity of monkey striate cortex: A parallel relationship between field size, scatter, and magnification factor. *Journal of Comparative Neurology*, 158:295-305, 1974.
- [8] C. Braccini, G. Gambardella, G. Sandini, V. Tagliasco. A model of the early stages of the human visual system: Functional and topological transformations performed in the peripheral visual field. *Biological Cybernetics*, 44:47-58, 1982.

- [9] A.S. Rojer, E.L. Schwartz. Design considerations for a space-variant visual sensor with complex-logarithmic geometry. In *Proceedings of the 10th International Conference on Pattern Recognition*, pages 278-285, 1990.
- [10] M.D. Levine. *Vision in Man and Machine*. McGraw-Hill, USA, 1985.
- [11] G. Sandini, V. Tagliasco. An anthropomorphic retina-like structure for scene analysis. *Computer Graphics and Image Processing*, 14:365-273, 1980.
- [12] E.L. Schwartz. Spatial mapping in the primate sensory perception: Analytic structure and relevance to perception. *Biological Cybernetics*, 25:181-194, 1977.
- [13] E.L. Schwartz. Anatomical and physiological correlates of visual computation from striate to infero-temporal cortex. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-14:257-271, 1984.
- [14] S.W. Wilson. On the retino-cortical mapping. *International Journal of Man-Machine Studies*, 18:361-389, 1983.
- [15] B.B. Bederson, R.S. Wallace, E.L. Schwartz. A miniaturized active vision system. In *Proceedings of the 1992 International Conference on Pattern Recognition*, pages 58-61, 1992.
- [16] M. Bolduc. *A Foveated Sensor for Robotic Vision*, Master's thesis, McGill University, Montréal, 1994.
- [17] R.F. Pierret. *Semiconductor Fundamentals*. Addison-Wesley Publishing Co., Reading, MA, 1988.
- [18] G.W. Neudeck, *The PN Junction Diode*. Addison-Wesley Publishing Co., Reading, MA, 1989.
- [19] S.M. Sze. *Physics of Semiconductor Devices*. Wiley, New York, NY, 1981.
- [20] D.K. Schroder. *Advanced MOS Devices*. Addison-Wesley Publishing Co., Reading, MA, 1987.
- [21] S.G. Chamberlain. Photosensitivity and scanning of silicon image detector arrays. *IEEE Journal of Solid-State Circuits*, SC-4:333-342, 1969.

- [22] N. Koike, I. Takemoto, K. Satoh, S. Hanamura, S. Nagahara, and M. Kubo. MOS area sensor: Part I-design consideration and performance of an n-p-n structure 484 x 384 element color MOS imager. *IEEE Transactions on Electron Devices*, ED-27:1676-1681, 1980.
- [23] D.F. Barbe. Imaging devices using the charge-coupled concept. *Proceedings of the IEEE*, 63:38-67, 1975.
- [24] R.H. Dyck and G.P. Weckler. Integrated arrays of silicon photodetectors for image sensing. *IEEE Transactions on Electron Devices*, ED-15:196-201, 1968.
- [25] A.B. Grebene. *Bipolar and MOS Analog Integrated Circuit Design*. John Wiley & Sons, Inc., New York, NY, 1984.
- [26] D. Brown, and A. Scott. *Design Rules and Process Parameters for the Northern Telecom CMOS4S process*. Canadian Microelectronics Corporation, Kingston, Ontario, 1990.
- [27] V. Ward, M.Syrzycki, and G. Chapman. CMOS photodetector with built-in light adaptation mechanism. *Microelectronics Journal*, 24:547-553, 1993.
- [28] R.F. Pierret. *Field Effect Devices*. Addison-Wesley Publishing Co., Reading, MA, 1983.
- [29] S.G. Chamberlain and J.P.Y. Lee. A novel wide dynamic range silicon photodetector and linear imaging array. *IEEE Journal of Solid-State Circuits*, SC-19:41-48, 1984.
- [30] T. Nakamura, K. Matsumoto, R. Hyuga and A. Yusa. A new MOS image sensor operating in a non-destructive readout mode. *Technical Digest - IEEE International Electron Devices Meeting*, pages 353-356, 1986.
- [31] J. Hyncek. BCMD—An improved photosite structure for high-density image sensors. *IEEE Transactions on Electron Devices*, ED-38:1011-1020, 1991.
- [32] K. Matsumoto, I Takayanagi, T. Nakamura, and R. Ohta. The operation mechanism of a charge modulation device (CMD) image sensor. *IEEE Transactions on Electron Devices*, ED-38:989-998, 1991.
- [33] G.P. Weckler. Operation of p-n junction photodetectors in a photon flux integrating mode. *IEEE Journal of Solid-State Circuits*, SC-2:65-73, 1967.

- [34] P.K. Weimer, G. Sadasiv, J.E. Meyer, L. Meray-horvath, and W.S. Pike. A self-scanned solid-state image sensor. *Proceedings of the IEEE*, 55:1591-1603, 1967.
- [35] P.J.W. Noble. Self-scanned image detector. *IEEE Transactions on Electron Devices*, ED-15:202-209, 1968.
- [36] A.M. Dhake. *Television Engineering (CCIR Systems-B Standards)*. McGraw-Hill Book Co., New York, NY, 1979.
- [37] R. Melen. The tradeoffs in monolithic image sensors: MOS vs. CCD. *Electronics*, 46:106-111, 1973.
- [38] J.D. Plummer, and J.D. Meindl. MOS electronics for a portable reading aid for the blind. *IEEE Journal of Solid-State Circuits*, SC-7:111-119, 1972.
- [39] N. Tanaka, T. Ohmi, Y. Nakamura, and S. Matsumoto. A low-noise Bi-CMOS linear image sensor with auto-focusing function. *IEEE Transactions on Electron Devices*, ED-36:39-45, 1989.
- [40] EG&G Reticon, *Image Sensing and Solid-State Camera Products*, 1994.
- [41] W.S. Boyle, and G.E. Smith. Charge coupled semiconductor devices. *Bell Systems Technical Journal*, 49:587-593, 1970.
- [42] Dalsa Inc., Waterloo, Ontario. *CCD Image Sensors and Cameras*, 1991.
- [43] E.R. Fossum. Active pixel sensors: are CCD's dinosaurs? *SPIE - Charge-Coupled Devices and Solid-State Optical Sensors III*, pages 2-14, volume 1900, 1993.
- [44] E.R. Fossum. Active pixel sensors challenge CCDs. *Laser Focus World*, 29:83-87, 1993.
- [45] Intel Corporation. *16-bit embedded controllers*, 1989.
- [46] N.C. Battersby. *Switched-current techniques for analogue sampled-data signal processing*, Phd. Thesis, University of London, 1993.
- [47] C.L. Keast, Private Communication, 1994.
- [48] C.L. Keast, and C.G. Sodini. A CCD/CMOS- based imager with integrated focal plane signal processing. *IEEE Journal of Solid-State Circuits*, SC-28:431-437, 1993.

- [49] E.R. Fossum. Charge-coupled computing for focal plane image preprocessing. *Optical Engineering*, 26:916-922, 1987.
- [50] N. Weste, and K. Eshraghian. *Principles of CMOS VLSI Design: A Systems Perspective*. Addison-Wesley Publishing Co., Reading, MA, 1985.
- [51] K. Chen, A. Astrom, and P.E. Danielsson. PASIC. A smart sensor for computer vision. *Proceedings of the 10<sup>th</sup> International Conference on Pattern Recognition*, pages 286-291, 1990.
- [52] M. Tremblay, M. d'Anjou, D. Poussart. Hexagonal sensor with embedded analog image processing for pattern recognition. *Proceedings of the IEEE 1993 Custom Integrated Circuits Conference*, pages 12.7.1-12.7.4.
- [53] D. Renshaw, P.B. Denyer, G. Wang, and M.Y. Lu, ASIC Vision. *Proceedings of the IEEE 1990 Custom Circuits Conference*, pages 7.3.1-7.3.4.
- [54] C.A. Mead, and M.A. Mahowald. A silicon model of early visual processing. *Neural Networks*, 1:91-97, 1988.
- [55] J. Nakamura, S.E. Kemeny, and E.R. Fossum. CMOS active pixel image sensor with simple floating gate pixels. *IEEE Transactions on Electron Devices*, ED-42:1693-1694, 1995.
- [56] T.M. Bernard, B.Y. Zavidique, and F.J. Devos. A programmable artificial retina. *IEEE Journal of Solid-State Circuits*, SC-28:789-798, 1993.
- [57] A.P. Marriott, Ph. Tsalides, P.J. Hicks. VLSI implementation of smart imaging system using two-dimensional cellular automata. *IEE Proceedings-G*, 138:582-586, 1991.
- [58] C.P. Chong, C.A.T. Salama, and K.C. Smith. An imager with built-in image-velocity computation capability. *IEEE Transactions on Circuits and Systems for Video Technology*, 2:306-312, 1992.
- [59] O. Yadid-Pecht, R. Ginosar, Y.S. Diamond. Random access photodiode array for intelligent image capture. *IEEE Transactions on Electron Devices*, ED-38:1772-1780, 1991.

- [60] M. Tremblay, D. Laurendeau, D. Poussart. High resolution smart image sensor with integrated parallel analog processing for multiresolution edge extraction. *Robotics and Autonomous Systems*, 11:231-242, 1993.
- [61] S.K. Mendis, S.E. Kemeny, and E.R. Fossum. CMOS active pixel image sensor. *IEEE Transactions on Electron Devices*, ED-41:452-453, 1994.
- [62] N. Ricquier, B. Dierickx. Pixel structure with logarithmic response for intelligent and flexible imager architectures. *Microelectronic Engineering*, 19:631-634, 1992.
- [63] P.B. Denyer, D. Renshaw, G. Wang, M.Y. Lu, S. Anderson. On-chip CMOS sensors for VLSI imaging systems. *IFIP Transactions A [Computer Science and Technology]*, A-1:157-166, 1992.
- [64] S. Anderson, W.H. Bruce, P.B. Denyer, D. Renshaw, and G. Wang. A single chip sensor and image processor for fingerprint verification. *Proceedings of the IEEE 1991 Custom Integrated Circuits Conference*, pages 12.1.1-12.1.4.
- [65] O. Vellacott. CMOS in camera. *IEE Review*, 40:111-114, 1994.
- [66] Integrated Vision Products, AB. Technical literature, 1993.
- [67] P.W. Fry, P.J. Noble, and R.J. Rycroft. Fixed-pattern noise in photomatrices. *IEEE Journal of Solid-State Circuits*, SC-5:250-254, 1970.
- [68] W.C. McColgin, J.P. Lavine, J. Kyan, D.N. Nicholas, and C.V. Stancampiano. Dark current quantization in CCD image sensors. *Technical Digest - IEEE International Electron Devices Meeting*, pages 113-116, 1992.
- [69] L. Jasterzebski, R. Soydan, G.W. Cullen, W.N. Henry, and S. Vecrumba. Silicon wafers for CCD imagers. *Journal of the Electrochemical Society*, 134:212-221, 1987.
- [70] M.H. White, D.R. Lampe, F.C. Blaha, and I.A. Mack. Characterization of surface channel CCD image arrays at low light levels. *IEEE Journal of Solid-State Circuits*, SC-9:1-12, 1974.
- [71] S.K. Mendis, S.E. Kemeny, R.C. Gee, B. Pain, Q. Kim, and E.R. Fossum. Progress in CMOS active pixel image sensors. *SPIE - Charge-Coupled Devices and Solid-State Optical Sensors IV*, pages 19-29, volume 2172, 1994.

- [72] R.R. Buss, S.C. Tanaka, G.P. Weckler. Principles of low-noise signal extraction from photodiode arrays. In P.G. Jespers, F. Van de Wiele, and M.H. White editors, *Solid State Imaging*, pages 561-603. Noordhoff, Leyden, 1976.
- [73] S. Ohba, M. Nakai, H. Ando, S. Hanamura, S. Shimada, K. Satoh, K. Takahashi, M. Kubo, T. Fujita. MOS area sensor: Part II-low-noise MOS area sensor with antiblooming photodiodes. *IEEE Transactions on Electron Devices*, ED-27:1682-1687, 1980.
- [74] Y. Nishida, J. Koike, H. Ohtake, M. Abe, and S. Yoshikawa. Design concept for a low-noise CCD image sensor based on subjective evaluation. *IEEE Transactions on Electron Devices*, ED-36:360-366, 1989.
- [75] C. Jansson, P. Ingelhart, C. Svensson, and R. Forchheimer. An addressable  $256 \times 256$  photodiode image sensor array with an 8-bit digital output. *Analog Integrated Circuits and Signal Processing*, 4:37-49, 1993.
- [76] N. Ricquier, and B. Dierickx. Random addressable CMOS image sensor for industrial applications. *Sensors and Actuators A - Physical*, 44:29-35, 1994.
- [77] Thompson-CSF, *The CCD Image Sensor*, 1992.
- [78] S. Ohba, M. Nakai, H. Ando, K. Takahashi, M. Masuda, I. Takemoto, and T. Fujita. Vertical smear noise model for an MOS-type color imager. *IEEE Transactions on Electron Devices*, ED-32:1407-1410, 1985.
- [79] H. Ando. MOS Imaging Devices. *Optoelectronics - Devices and Technologies*, 6:321-332, 1991.
- [80] A. Gruss, R. Carley, and T. Kanade. Integrated sensor and range-finding analog signal processor. *IEEE Journal of Solid-State Circuits*, SC-26:184-191, 1991.
- [81] B.W. Kernighan, D.M. Ritchie. *The C Programming Language*. Prentice Hall, Englewood Cliffs, NJ, second edition, 1988.
- [82] K.R. Lakshmikumar, R.A. Hadaway, and M.A. Copeland. Characterization and modeling of mismatch in MOS transistors for precision analog design. *IEEE Journal of Solid-State Circuits*, SC-21:1057-1066, 1986.
- [83] E. Hecht. *Optics*. Addison-Wesley Publishing Co., Reading, MA, second edition, 1987.

- [84] The MathWorks, Inc., *MATLAB Reference Guide*, 1992.
- [85] H. Yamamoto, Y. Yeshurun, M.D. Levine. An active foveated vision system: Attentional mechanisms and scan path convergence measures. *CVGIP:Image Understanding*, accepted for publication.
- [86] P.B. Denyer, D. Renshaw, W. Guoyu, L. Mingying. CMOS image sensors for multimedia applications. *Proceedings of the IEEE 1993 Custom Integrated Circuits Conference*, pages 11.5.1-11.5.4.
- [87] D.V. Plant, B. Robertson, H.S. Hinton, W.M. Robertson, G.C. Boisset, N.H. Kim, Y.S. Liu, M.R. Otazo, D.K. Rolston, A.Z. Shang. An optical backplane demonstrator system based on FET-SEED smart pixel arrays and diffractive lenslet arrays. *IEEE Photonics Technology Letters*, 7:1057-1059, 1995.
- [88] F.B. McCormick, T.J. Cloonan, A.L. Lentine, J.M. Sasian, R.L. Morrison, M.G. Beckman, S.L. Walker, M.J. Wojcik, S.J. Hinterlong, R.J. Crisci, R.A. Novotny, H.S. Hinton. Five-stage free-space optical switching network with field-effect transistor self-electro-optic-effect-device smart-pixel arrays. *Applied Optics*, 33:1601-1618, 1994.