A Wide Dynamic Range CMOS Imager With Parallel On-Chip Analog-to-Digital Conversion

by

Steven John Decker

B.S., The Ohio State University (1988)
S.M., Massachusetts Institute of Technology (1991)

Submitted to the Department of Electrical Engineering and Computer Science
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1997

Signature of Author .....
Department of Electrical Engineering and Computer Science
September 3, 1997

Certified by .....
Charles G. Sodini
Professor of Electrical Engineering and Computer Science
Thesis Supervisor

Accepted by .....
Arthur C. Smith
Chair, Department Committee on Graduate Students

A Wide Dynamic Range CMOS Imager With Parallel On-Chip Analog-to-Digital Conversion

by

Steven John Decker

Submitted to the Department of Electrical Engineering and Computer Science
on September 5, 1997, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

Abstract

The imager is the front-end of any machine vision system. Traditional machine vision systems have used CCD video cameras with a single analog output channel, which results in several limitations. The single output channel is generally designed for human vision, around 30 frames/sec, which is much slower than desirable for some machine vision tasks. The imager dynamic range is low compared to the scene dynamic range, which can be as much as 100,000:1 for automotive applications. The analog output format is incompatible with digital processing chips, requiring additional chip(s) for correlated double-sampling, gain control, and analog-to-digital conversion.

This thesis describes the theory, design, and characterization of a prototype 256 $\times$ 256 CMOS active pixel sensor (APS) array which addresses these problems. Array output is column-parallel, allowing a high frame rate. Dynamic range is increased using a technique which varies the lateral overflow gate voltage over the integration period. The compression curve is digitally controlled and user-adjustable. Fixed-pattern noise (FPN) is partially removed by a switched-capacitor circuit. Imager output is digitized on-chip by cyclic analog-to-digital converters.

The prototype imager was manufactured in a 0.8 μm CMOS process with linear (poly - n^+) capacitors. The imager was characterized at 30 frames/sec, although operation was verified for frame rates between 30 and 400 frames/sec. The saturation level was 1.69 V, and random noise was 0.57 mV (1 σ), producing a dynamic range of 3000 for the linear imaging mode. The wide dynamic range imaging technique increases dynamic range by 20$\times$ with good image quality, and by approximately 160$\times$ with reduced image quality. The static power dissipation was 52 mW. Pixel FPN for a dark frame was 3.9 mV (1 σ), and column FPN was 1 mV (1 σ).

Thesis Supervisor: Charles G. Sodini
Title: Professor of Electrical Engineering and Computer Science

Dedication

This thesis is dedicated to my parents, Ronald and Shirley Decker.

Acknowledgements

Funding for this work was provided through the National Science Foundation and the Advanced Projects Research Administration under contract MIP-9423221.

I wish to thank my advisor, Professor Charles Sodini, for suggesting the thesis topic. His technical guidance, positive attitude, and ability to see the fundamental issues have been extremely valuable. I deeply appreciate his steady support for the project, and for me, over the past several years.

Kevin Brehmer hosted me for a month at National Semiconductor Corporation in Santa Clara, and provided the chip fabrication. His hospitality and technical input have been invaluable. I would like to thank my thesis readers, Dan McGrath and Hae-Seung (Harry) Lee, for freely sharing their knowledge and experience, and for taking the time to read and comment on the thesis. Professor John Wyatt provided valuable insight into the operation of the wide dynamic range pixel, and invested enormous effort in forming and leading the MIT Vision Group. Ichiro Masaki has continued the effort, and ensured continued funding so that this project could be completed.

Jen Lloyd and Jeff Gealow invested an enormous amount of time in bringing up the CAD tools used in this project. Their effort, and their willingness to help with all types of computer trouble, are deeply appreciated. I would like to thank them and my other colleagues, including Joe Lutsky, Chris Umminger, Ignacio McQuirk, Michael Ashburn, Dave Martin, and Andy Karanicolas, for providing insights, discussion, and friendship. My UROP, Steve Kim, is also owed a debt of gratitude for his efforts in producing a working camera.

Finally, I would like to thank my family for their constant understanding and support.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Solid-state imaging is now an integral part of modern life. For example, camcorders, digital still cameras, surveillance cameras, electronic news gathering, and machine vision systems use solid-state imagers frequently or exclusively. These applications have very different requirements for resolution, frame rate, output signal format, and so on, which impose different design challenges for the imager. In this thesis, an imager for machine vision is developed.

A machine vision system, shown in Fig. 1-1, consists of an imager and either an analog signal processor or an analog-to-digital converter (ADC) and digital processor. The imager is typically a CCD imager with a single analog output, a dynamic range between perhaps 100:1 and 1000:1, and a low (~30 Hz) frame rate. The imager is a limitation for applications which require high frame rate, such as assembly line inspection systems, or applications which require high dynamic range, such as automotive vision systems. In addition, it is desireable to integrate the system to reduce its power, size, and cost.

The MIT Machine Vision Group [1] has been exploring systems which use analog signal processing and varying levels of integration. Several chips designed in the course of this project

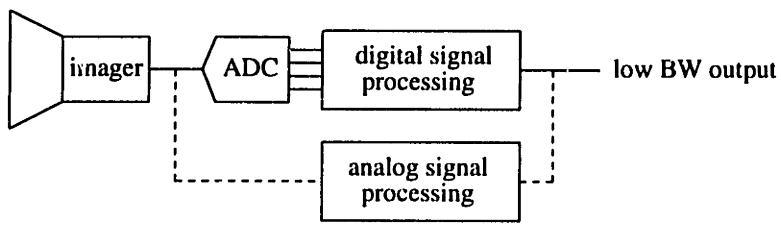


Figure 1-1: Block diagram of a machine vision system. A given system will use either analog or digital signal processing.

have integrated the imager with an analog processor using full 2-D connectivity [2, 3, 4]. The benefits include a wide data path between imager and processor, elimination of the ADC, and reduced system size. Impressive results have been achieved for certain machine vision tasks; for example, one chip is capable of locating the center of mass and orientation of an object at 5000 fps [2]. One drawback of this architecture is the conflict between making a good imager and making a highly functional processor. As the processor area increases, the light-collecting area of the pixel decreases, decreasing quantum efficiency. Since each pixel is generally placed next to a processing element, the pixel array is spread out over the entire chip, requiring an expensive large-format lens. Another drawback is that the output from the imager can only be used by the processor on that chip.

Quantum efficiency and pixel array size are improved by physically separating the pixel array from the processing circuits. Integrated systems employing this architecture include a focus-of-expansion chip [5] and a 2-D discrete cosine transform chip [6]. In both cases, the analog processor has limited flexibility, and pixel array output is available only to the on-chip processor.

In some systems, the analog processor is on a different chip from the pixel array [7, 8]. The advantages are that more area is available for both pixel array and processor, and pixel array output is available for multiple processors. The disadvantage is decreased integration. A programmable analog processing element tends to be large; a recent design required 700 μm $\times$ 270 μm [8] per processing element in a 0.8 μm CMOS process.

Digital processing allows higher accuracy, and larger processing arrays on a chip [9]. This requires analog-to-digital conversion, which could be performed on a separate chip, the digital processor chip, or the imaging chip. Using a separate chip decreases system integration. Column-parallel conversion on the processing chip requires multiple analog channels between the imager and processor, and maintaining low crosstalk between adjacent channels is a potential drawback. Conversion on the imager chip mostly avoids the crosstalk problem, and is the more common choice. For example, Paul [10] has described a charge-based converter which produces serial digital output from a CCD pixel array. Column-parallel approaches have also been described [11, 12] using single-slope converters and sigma-delta converters. A focal-plane, column-parallel cyclic converter was chosen for the present work.

In addition to on-chip data conversion, the usefulness of this imager is improved by the use of a technique which increases the dynamic range. Low dynamic range is seen in the image as clipping in the dark and/or bright regions. The wide dynamic range technique used in this thesis has several advantages over other previously described schemes [13, 14, 15, 16, 17, 18, 19, 20, 21].

It is capable of approximating any compressive imaging characteristic. The compression curve (Q_{int} vs. Irradiance) is user-adjustable, and can be effectively turned off if desired. Pixel area is the same as for an ordinary active pixel sensor (APS) pixel because it uses a transistor which doubles as the reset transistor.

The main contribution of this thesis is a practical method of employing the wide dynamic range pixel in an area array. The pixel itself has been previously described in a CCD version, but no method has been described for generating a uniform lateral overflow gate waveform without requiring charge storage in the array. A further contribution of this thesis is the fully-differential signal processing chain for correlated double-sampling and column-parallel analog-to-digital conversion.

This thesis describes many advancements in the state of APS design. Although the dynamic range enhancement technique is not original with this thesis, it is shown that the continuous-time barrier described by previous authors can be replaced by a stepped barrier function with good results. A general mathematical relationship between the compression curve and barrier function is derived. A practical pixel array architecture which avoids dynamic range degradation due to optical crosstalk is developed.

The thesis is organized as follows. Chapter 2 covers background material on CCD and CMOS imaging, previously described techniques for dynamic range enhancement, and techniques for focal plane analog-to-digital conversion. Chapter 3 discusses the continuous-time and stepped-time theory behind the dynamic range enhancement technique. Chapter 4 covers the design of the prototype imager. Chapter 5 presents test data from the prototype. Chapter 6 summarizes the results and presents suggestions for future work and enhancements.

Chapter 2

Background

A pixel array transduces light to an electrical quantity (charge, current, or voltage) over a two-dimensional scene. Nearly all visible-light solid-state imagers are based on the principle of carrier photogeneration. Photons entering a semiconductor generate electron-hole pairs if the photon energy exceeds the bandgap energy. An electric field separates the holes and electrons; typically, the electrons are collected in a potential well and the holes are discarded. The major differences between imager types are related to how the electrons are converted to an output signal and how the array is scanned out.

Two technologies for imaging arrays were considered for this project. The charge-coupled device (CCD) array has been the dominant solid-state imager since its invention in 1970 [22]. CCD arrays make excellent imagers. Read noise as low as 3 e^- per pixel has been achieved [23]. Pixel sizes as small as 2.4 μm $\times$ 2.4 μm can be achieved in currently available technologies [24]. Fixed-pattern noise (FPN) in the dark is so low that it is seldom reported; FPN in a recent CCD imager was reported to be 10 e^- while using a dark current reduction technique [25].

Several MOS imagers were built before 1970 [26] and MOS imagers have recently experienced a resurgance of interest. Compared to CCD imagers, MOS imagers generally have higher read noise, higher fixed-pattern noise, and larger pixels. However, they can be built in readily available CMOS processes, and thus they can take advantage of the constant advances in minimum feature size. Digital and analog circuitry are more easily integrated with the imager. The CMOS imager has other advantages for wide dynamic range imaging which will be explained later.

This chapter begins by defining several imager parameters. CCD imaging arrays are discussed next. The principle of CCD operation is explained. Photogate and photodiode CCD

pixels are described, and the common pixel array architectures are reviewed. MOS pixels and architectures are described. CCD and CMOS technologies are compared for imaging applications. Several techniques for increasing the effective pixel dynamic range are discussed. Finally, several analog-to-digital converter architectures will be compared for focal-plane data conversion.

2.1 Imager Performance Characterization

Imagers are commonly characterized by quantum efficiency, conversion gain, random noise, FPN, lag, smear, saturation level, anti-blooming, and dynamic range. These terms will be defined below.

2.1.1 Quantum Efficiency

Quantum efficiency is the probability that a photon incident on the pixel creates a hole or electron that is collected in the pixel's potential well. Quantum efficiency depends on several parameters, including the reflection coefficient of the optical stack on the pixel, the fraction of the pixel area which is light-shielded, the fraction of the pixel area where photocharge goes elsewhere than the photosite, the photon wavelength, the bulk and surface carrier lifetimes, the depth of the potential well, and the volume and shape of the depletion regions. It is useful to think of quantum efficiency as the product of optical aperture, which is the fraction of pixel area which is optically active, optical stack transmission, which is the fraction of photons incident on optically active areas which reach the silicon surface, and collection efficiency, which is the fraction of photogenerated electrons that get collected.

2.1.2 Conversion Gain

The image data starts as charge and is usually converted to voltage. For an imager with a linear charge-to-voltage characteristic, the ratio of output signal to collected charge is a constant, referred to as the conversion gain. It is usually expressed in units of $\mu V/e^-$ and is equal to the reciprocal of the charge sense capacitance, times any system gain.

2.1.3 Random Noise

An imager with a constant scene should produce identical output from frame to frame. In practice, the output from a given pixel will vary over time due to thermal noise, charge trapping,

and $1/f$ noise in the devices which comprise the imager. Photonic shot noise is usually not included in this quantity, although this also contributes to noise at the output. Random noise is typically stated in terms of input-referred equivalent electrons, *i.e.*, the root mean square (rms) output voltage noise divided by the conversion gain.

2.1.4 Saturation Level

Pixel saturation occurs when the pixel stores the maximum amount of charge. If more charge is added, it simply spills into the substrate or adjacent pixels. It is usually expressed in terms of electrons, although it is sometimes referred to the output as a voltage. It can also be stated in terms of the illumination required to produce the saturation charge. Depending on design, the output circuit can saturate before the pixel saturates. In this case, the saturation level is set by the output circuit.

2.1.5 Fixed-Pattern Noise

Fixed-pattern noise (FPN) is the fixed (constant in time) variation between pixel outputs under spatially uniform illumination. Fixed-pattern noise is typically due to random or mask-induced mismatches in device parameters such as threshold voltage, trap density, parasitic capacitance and transistor β. FPN is usually a function of illumination, and can be written as the sum of a gain term and an offset term for an imager with a linear response characteristic. Offset FPN is constant over illumination, and gain FPN is proportional to illumination.

FPN consists of components that describe variation between columns, and variation between pixels in a single column. Column FPN is the standard deviation of the column-average pixel output values in a time-average, uniformly illuminated frame.

$$\text{FPN (column)} = \sqrt{\frac{\Sigma_j(\overline{P}_j - \overline{P})^2}{j-1}}, \tag{2.1}$$

where $\overline{P}_j$ is the average pixel value in column j and $\overline{P}$ is the average pixel value in the frame. Since a column FPN calculation requires multiple columns, $j > 1$. Pixel FPN is the standard deviation of pixel output values after column FPN has been removed. In order to calculate pixel FPN, multiple pixels are required, so $i \cdot j > 1$.

$$\text{FPN (pixel)} = \sqrt{\frac{\Sigma_{i,j}(P_{i,j} - \overline{P}_j)^2}{i \cdot j - 1}}. \tag{2.2}$$

2.1.6 Image Lag

Image lag refers to the persistance of one frame in successive frames. An imager demonstrates lag when, for example, a strong illuminant is suddenly turned off but the imager output does not immediately return to the black level. Lag is usually caused by incomplete charge reset. It is informally characterized by whether or not it is visible on a video display.

2.1.7 Smear

In a CCD imager, pixel charge is clocked out through a shift register. A charge packet passing by a region of high illumination accumulates additional charge due to optical crosstalk, creating vertical stripes in the reconstructed image. Smear is characterized by the charge accumulated by a potential well passing by a column of pixels 10% the total column height, with the pixels illuminated by the saturation illumination. For example, assume a 1000 $\times$ 1000 pixel interline CCD imager. The imager is completely in the dark, except for 100 rows which are illuminated with the saturation illumination. An empty potential well in the vertical shift register, starting at the top of the shift register, will acquire charge Q_{smear} by the time it reaches the bottom. Smear is expressed in dB as the ratio of the saturation charge to Q_{smear}.

Charge readout CMOS imagers can exhibit the same behavior. A column line can easily collect stray charge from a highly illuminated pixel, corrupting the readout of all pixels in that column.

2.1.8 Crosstalk

Crosstalk is any contamination of one pixel's signal by another pixel's signal. Smear is an example of optical crosstalk. Examples of electrical crosstalk include ground bounce and capacitive coupling.

2.1.9 Anti-Blooming

Blooming is the spread of charge from *saturated* pixels into surrounding potential wells. It differs from optical crosstalk in that all photogenerated charge, not just a small fraction, spreads out into the surrounding pixels/shift registers. Most pixels contain some means for shunting excess charge to a drain in order to preserve the information in nearby non-saturated pixels. Anti-blooming characterizes the effectiveness of the technique used to drain excess charge. It is expressed as the ratio of the illumination required to produce blooming to the illumination

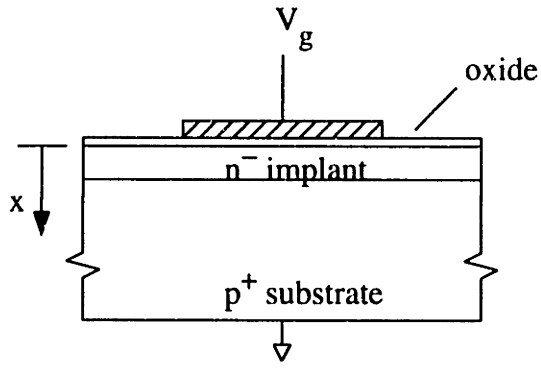


Figure 2-1: MOS capacitor structure.

required to saturate the pixel. For example, a pixel with no anti-blooming protection would have 1× anti-blooming.

2.2 CCD Imagers

The CCD is the most common solid-state imager. The basis of this imager is the ability to completely transfer charge between adjacent MOS capacitors by applying overlapping clocks to their gates. The MOS capacitor with buried-channel implant is shown in Fig. 2-1. The buried-channel implant is a lightly n-doped region that can be fully depleted; it keeps the charge packet away from the SiO_2 interface, reducing charge loss in the interface traps and creating fringing fields to allow higher transfer speed. Figure 2-2(a) shows the potential profile for several different gate voltages with a fully depleted buried-channel implant. In practice, the gate voltage is set to one of two voltages: high to store charge and low to remove all charge. A third level is sometimes used to enable charge transfer through an alternate path. Figure 2-2(b) shows the potential profile for a fixed (high) gate voltage and various amounts of stored charge [27]. Free charge is indicated by a flat region at the bottom of the potential profile. The potential well is full when charge spills into adjacent pixels or through an overflow drain, or when the flat region extends to the silicon surface, since the MOS capacitor becomes a surface-channel device at that point.

Charge is transfered between adjacent MOS capacitors by applying overlapping clocks to their gates. The concept is demonstrated by Fig. 2-3, which shows two adjacent MOS gates and their corresponding potential energy diagrams. The potential energy diagrams are shown using the fluid model of the CCD, where the y-axis is electron potential energy, and charge is thought of as a fluid that flows to the state of lowest potential energy. In Fig. 2-3(a), the charge

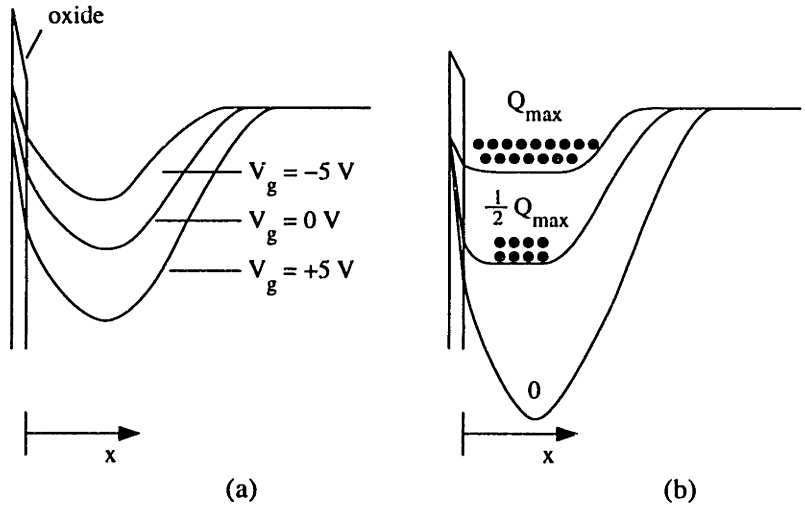


Figure 2-2: (a) Buried-channel potential profiles for zero charge. (b) Buried-channel profile for fixed gate voltage and varying charge.

is entirely under gate 1. In (b), the voltage on gate 2 is high so that charge is shared evenly between the two gates. In (c), the voltage on gate 1 goes low, transfering all charge under gate 2. Note that as voltage on a gate increases, the bottom of its potential energy well is lowered. This happens because potential energy diagrams are usually drawn for electrons, which have negative charge.

2.2.1 CCD Pixels

The CCD pixel is an integrating type, *i.e.*, charge is accumulated over the integration time and sent to the output circuit at the end of the integration period. In its simplest form, a pixel is just a potential well and a transfer gate. The potential well can be created by either a photogate or a photodiode.

A photogate pixel, shown in Fig. 2-4(a), uses a MOS capacitor to create a potential well. During integration, the transfer gate (TX) voltage is low and the photogate (PG) voltage is high. Charge is transfered out of the pixel by raising the transfer gate and lowering the photogate. The vertical shift register phase $\phi 1$ is high, creating a potential well to receive the charge. Photogate pixels tend to have a high saturation level due to the large gate capacitance. They also tend to have poor quantum efficiency, especially in the shorter wavelengths, due to absorption in the polysilicon gate.

A simple photodiode pixel is shown in Fig. 2-4(b). Operation is similar to the photogate

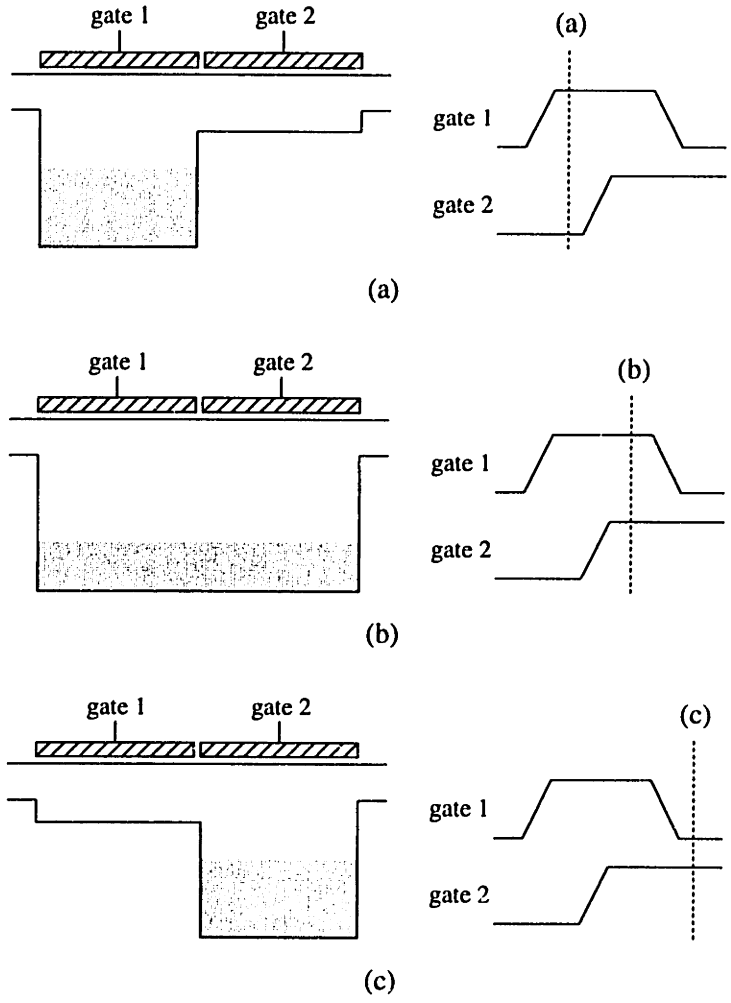


Figure 2-3: (a) All charge under gate 1. (b) Charge evenly divided between gate 1 and gate 2. (c) All charge under gate 2.

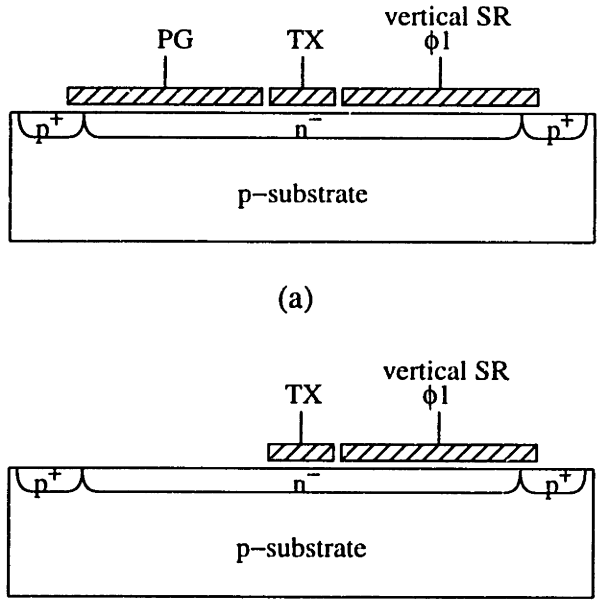


Figure 2-4: (a) Photogate CCD pixel. (b) Photodiode CCD pixel.

pixel. During integration, the transfer gate isolates the charge packet from the vertical shift register. At the end of the integration period, TX and $\phi 1$ go high, skimming the integrated charge in the photodiode into the vertical shift register. Then TX goes low, starting another integration period. Readout of a photodiode pixel is relatively slow since transfer of the last few electrons takes place solely by diffusion. Quantum efficiency will be relatively high since no photogate blocks the short wavelengths.

2.2.2 CCD Architectures

Charge packets in the pixels must be transfered to the output circuit(s). The unique feature of CCD's is that this transfer occurs through a sequence of complete transfers between adjacent MOS capacitors. The details of the transfer process are specified by the architecture. Common CCD area imager architectures include the frame transfer (FT), interline transfer (IT), and frame interline transfer (FIT) designs [28].

A frame transfer array is shown in Fig. 2-5. At the end of the integration period, charge is rapidly transfered directly through the array to a light-shielded storage area. The frame is then read out from the storage area. Smear is a problem for this architecture, and the transfer speed is limited due to the large length of a pixel compared to the length of the transfer gate in

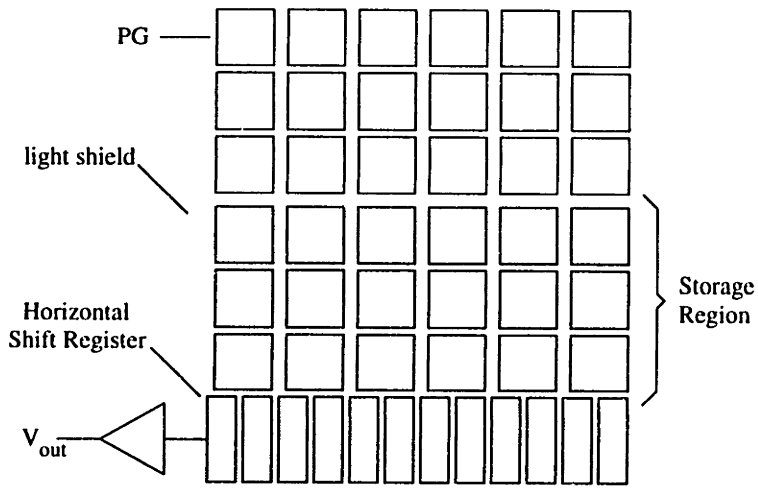


Figure 2-5: Frame transfer (FT) CCD array.

a typical CCD shift register. Optical aperture is excellent, approaching 100% in some imagers. This is somewhat misleading, however, as the channel stops and polysilicon photogates have poor optical response. A storage area, roughly equal in size to the imaging area, is required to store the frame until it can be read out.

An interline transfer (IT) array is shown in Fig. 2-6. A vertical shift register is placed between each column of pixels. At the end of the integration period, the transfer gate (TG) in the pixel and $\phi 1$ go high. Pixel charge spreads under those gates. As the transfer gate and photogate go low, the remainder of the charge is transfered under the $\phi 1$ gates. As soon as TG goes low, the pixel can start another integration period. Meanwhile, the charge packets in the vertical shift register are shifted down towards the horizontal shift register. When a new set of packets is shifted into the horizontal shift register, the vertical shift register halts until all of the charge in the horizontal shift register is transfered to the output circuit, where the packets are converted into voltages. The shift registers are usually timed so that it takes exactly one integration period to clock out all of the pixels in the array, so that by the time one frame has been read out, the next frame is ready to be read out.

The IT array suffers somewhat in optical aperture since the vertical shift register is light shielded, and is therefore optically inactive. However, smear is low compared to a frame transfer array operated at the same frame rate, since the charge is not shifted through optically active regions.

The FIT architecture combines the IT and FT architectures, as shown in Fig. 2-7. The advantages are the photodiode capability and very low smear. The disadvantages are reduced

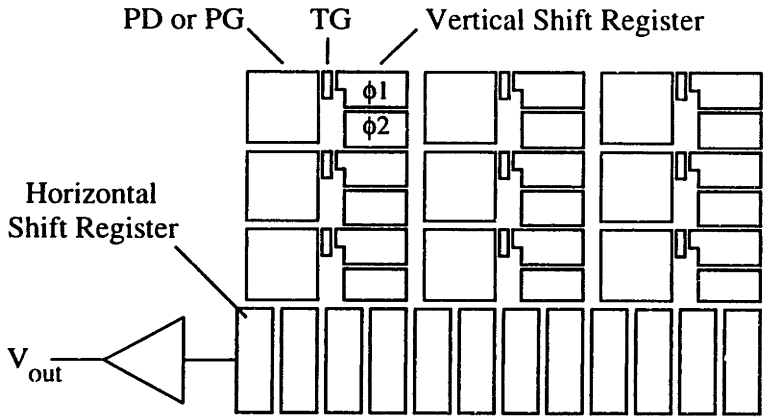


Figure 2-6: Interline transfer (IT) CCD array.

optical aperture due to the vertical shift register and an approximate doubling of the required silicon area because of the storage region.

2.2.3 CCD Performance

Quantum Efficiency

Optical aperture is strongly dependent on architecture type. An FT imager can have nearly 100% optical aperture. An IT or FIT imager has lower optical aperture since the shielded interline shift registers block at least some of the incident light. Shift register area is a largely a matter of design, subject to design rule limitations. For example, consider a photogate IT imager which uses the same buried-channel implant for integrating well and four-phase shift register. If it is necessary for the shift register to handle the full pixel capacity, the shift register area must be twice the photogate area, limiting optical aperture to 33%. If a photodiode is used, the integrating well capacitance decreases, allowing a small shift register and higher optical aperture. If pixel capacity is limited by a lateral overflow gate or vertical anti-blooming, the shift register can again be reduced. Shift register capacity can be increased by increasing the buried channel implant. In any case, the shift register can be no smaller than that allowed by minimum design rules.

Collection efficiency is determined by pixel geometry and doping profile. CCD pixels usually have carefully tailored doping profiles which can optimize collection efficiency given other design constraints, such as vertical anti-blooming or surface pinning [29].

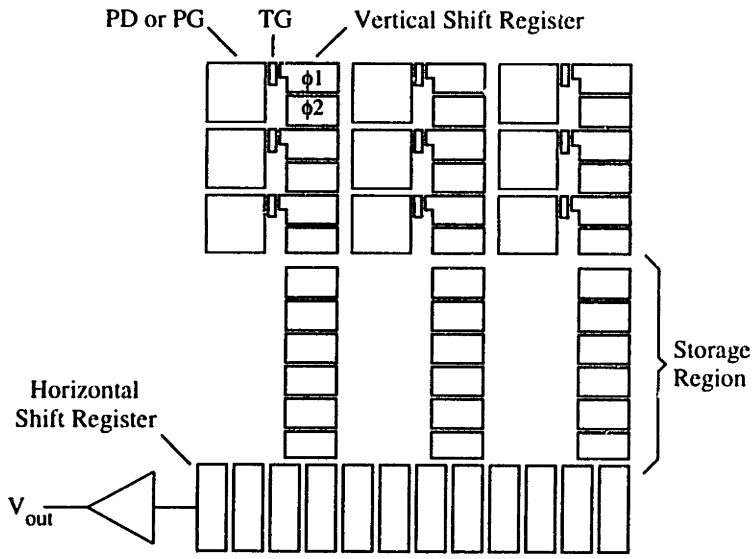


Figure 2-7: Frame interline transfer (FIT) CCD array.

Random Noise

Random noise includes photonic noise, dark current noise, charge transfer noise, switch noise, and amplifier noise. These components are added in quadrature to find the total random noise.

Photonic and dark current noise refer to the shot noise of the photocurrent and dark current, respectively. The noise for these sources is

$$N_{e,shot,rms} = \sqrt{\frac{1}{q}(Q_s + Q_d)}, \tag{2.3}$$

where Q_s is the signal charge and Q_d is the dark charge. Charge transfer noise [30] is due to a similar shot noise process; when a charge packet is transfered between potential wells, each electron has a probability ϵ of making the transfer. The total noise in the charge transfer is

$$N_{e,ct,rms} = \sqrt{2(1-\epsilon)N_s N_t}, \tag{2.4}$$

where N_s is the number of signal electrons and N_t is the number of transfers. Switch noise is caused by resetting the output diffusion. It has magnitude

$$N_{e,switch,rms} = \sqrt{\frac{kT}{q^2}C_{sense}}, \tag{2.5}$$

where C_{sense} is the capacitance of the output diffusion. The source-follower amplifier has input-referred $1/f$ and thermal noise. Switch noise and amplifier $1/f$ noise can be mostly removed

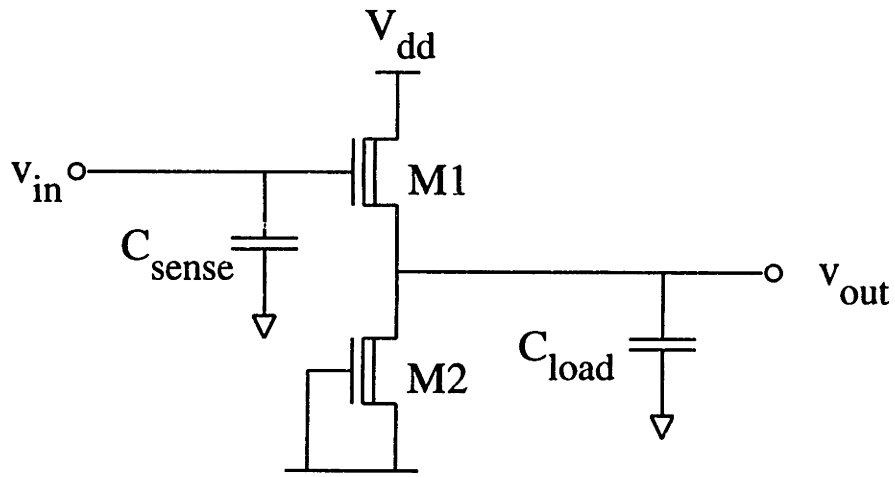


Figure 2-8: Output source follower for CCD imager.

using correlated double-sampling (CDS), but the source follower thermal noise is actually doubled by CDS. Source-follower thermal noise can be estimated from Fig. 2-8. Only one stage is shown, although in some cases, multiple source-follower stages are used. The thermal noise at the output is

$$N_{e,thermal} = \frac{C_{eff}}{qG_{sf}}\sqrt{\frac{8}{3}\frac{kT}{G_{out}^2}(g_{m1}+g_{m2})\Delta f} \tag{2.6}$$

$$= \frac{C_{eff}}{qG_{sf}}\sqrt{\frac{8}{3}\frac{kT}{G_{out}^2}(g_{m1}+g_{m2})\frac{G_{out}}{2\pi C_{load}}\frac{\pi}{2}} \tag{2.7}$$

$$= \frac{C_{eff}}{qG_{sf}}\sqrt{\frac{2}{3}\frac{kT}{G_{out}C_{load}}(g_{m1}+g_{m2})} \tag{2.8}$$

$$= \frac{C_{eff}}{qG_{sf}}\sqrt{\frac{2}{3}\frac{kT}{(g_{mb1}+g_{m1}\frac{C_{sense}}{C_{sense}+C_{gs1}})(C_{load}+\frac{C_{sense}C_{gs1}}{C_{sense}+C_{gs1}})}(g_{m1}+g_{m2})} \tag{2.9}$$

$$\approx \frac{C_{eff}}{q}\sqrt{\frac{kT}{C_{load}}\frac{g_{m1}+g_{m2}}{g_{mb1}+g_{m1}\frac{C_{sense}}{C_{sense}+C_{gs1}}}}, \tag{2.10}$$

where C_{eff} is given by

$$C_{eff} = C_{sense} + C_{gs1}(1 - G_{sf}), \tag{2.11}$$

and G_{sf} is the source follower gain.

The sum of the signal-independent noise sources is called read noise. Read noise varies from about 3 e^- to 200 e^- [23]. Shot noise has a square-root dependence on signal; above a certain irradiance, random noise is dominated by photon shot noise (charge transfer noise also

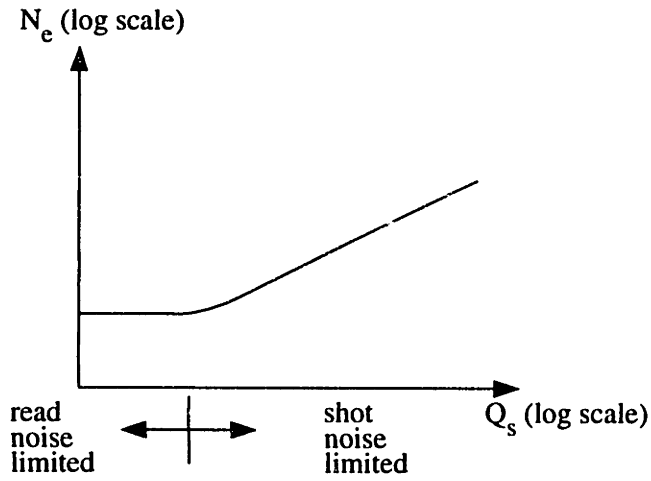


Figure 2-9: Signal dependence of random noise in a CCD imaging array.

has a square-root dependence on signal, but should be much lower than the photonic noise). Figure 2-9 shows the general noise dependence on signal level.

Saturation Level

The pixel saturation level is proportional to the integrating well capacitance, size, and voltage swing. The saturation level for a pixel in a CCD imager is usually in the range of 1000 to 3000 $e^-/\mu m^2$.

FPN

Fixed-pattern noise is usually negligible for CCD imagers, except for defective pixels. Defective pixels are defined as those having gain and/or offset FPN exceeding specified limits. CCD imagers are sorted into different grades, depending on the maximum number and type of defects allowed in specified regions of the array (see, for example, [31]). Typically, no more than a handful of defective pixels are tolerated in any grade.

Offset FPN occurs through pixel dark current I_d. Gain FPN arises from variation in pixel optical aperture A_{opt}, or any source of quantum efficiency mismatch. Gain FPN can become a problem as the aperture size becomes comparable to the polysilicon grain size. The net FPN is

$$\text{FPN} = T_{int}\left(\Delta I_d + I_s\frac{\Delta A_{opt}}{A_{opt}}\right), \tag{2.12}$$

where I_s is the pixel photogeneration current.

Table 2-1: Comparison of CCD pixel specifications.

Pixel Size (μm^2)	Q_{sat} (e^-)	$e^-/\mu m^2$	Smear (dB)	Architecture	Reference
6.0 × 12.6	70 000	930	80	FT	bosiers94a[32]
7.3 × 7.3	80 000	1500	77	IT	nichols92[33]
9.6 × 5.4	100 000	1900	140	FIT	yamashita94[34]
5.1 × 5.1	60 000	2300	75	FT	bosiers94b[35]
12 × 12	> 300 000	2100	NA	TDI	agwani95[36]

Image Lag

Photogate CCD's usually have no visible image lag, since pixel charge is fully transfered into the vertical shift register. Possible sources include bulk trapping and insufficient readout time. The most problematic bulk traps are those with time constants greater than the pixel readout time but less than the integration time. These traps fill during bright frames but do not discharge during readout. In a subsequent dark frame, the traps discharge without being refilled, adding charge which is indistinguishable from photogenerated charge. The bulk trap density is generally low enough that this is not a significant problem. Insufficient readout time causes lag by turning off the transfer gate before all integrated charge has been removed. This is seldom a problem because the time constants for charge transfer are usually quite small, and transfer time is relatively unconstrained.

Lag can be a significant problem in a photodiode CCD. Because charge transfer occurs by subthreshold conduction through a MOS transistor, the photodiode is never completely reset. Any residual charge in one frame appears in the following frame.

Smear

Smear is determined by the doping profiles of the integrating well and shift register, the proximity of the shift register to the aperture, and the configuration of the light shield over the vertical shift register. Smear is typically 70 - 80 dB (3.5 - 4 orders of magnitude) in an IT imager, although it can be much better in an FIT architecture. Table 2-1 lists the smear and saturation level for several recent CCD imagers.

Crosstalk

The CCD imager is robust against electrical crosstalk since voltage fluctuations in the substrate do not affect the number of carriers stored in a potential well. The vertical overflow drain (VOD), a common structure in CCD's, produces a potential barrier which inhibits injected

electrons from entering a pixel. Also, the epi-substrate interface has a high defect concentration, which tends to recombine injected electrons.

Optical crosstalk results from light entering one pixel which adds charge to another pixel's charge packet. One form is pixel-to-pixel crosstalk, which is present in all pixel arrays. This is relatively benign since it seldom occurs between pixels with widely varying intensities. Smear is a more serious form, as it limits the scene dynamic range, especially for FT and IT architectures. Smear can be greatly reduced by an FIT architecture, but this requires approximate doubling of the silicon area.

2.3 MOS Imagers

An MOS imager is any pixel array which uses MOSFET transistors to convey the signal from the pixels to the output circuit(s). Signal transfer occurs over a wire, and the signal can be charge, current, or voltage. Pixel output can be either continuous-time or integrated.

2.3.1 MOS Array Architecture

A typical MOS array architecture is shown in Fig. 2-10. The MOS array differs from the CCD array in that the column line can only carry one signal at a time, and must therefore be time multiplexed between all the pixels in a column. This requires a row select device in the pixel; when a row is read out, all other rows are disconnected from the column line. The outputs from all pixels in a row become available simultaneously. If only one output circuit is available for the entire array, the output columns must be multiplexed to one output.

The transduced signal is initially charge. MOS pixels can be classified according to the form of the signal between the pixel and the output circuit(s). The usual choices are charge, current, and voltage. The output circuit converts that signal to either an analog signal (normally a voltage) or a bit stream, and drives the signal off-chip.

2.3.2 Charge Output Pixel

Figure 2-11 shows a plan view and cross-section of a charge readout pixel. A photogate implementation is also possible; it would give a higher saturation charge but lower quantum efficiency. Figure 2-12 shows the equivalent circuit and timing diagram for a pixel and output circuit. The crosstalk current source represents the optical crosstalk current from other pixels on the same column line. The line capacitance is due to parasitic capacitances between the column

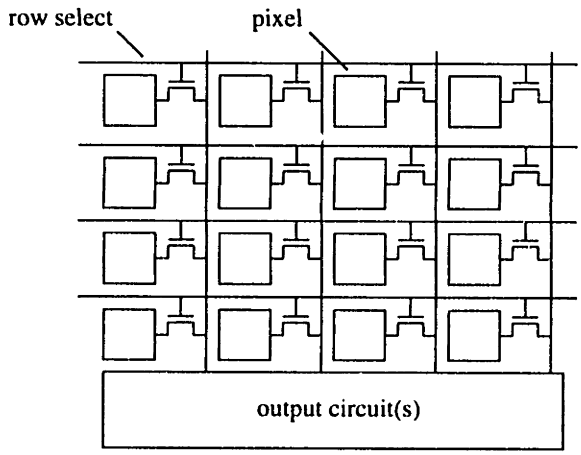


Figure 2-10: MOS imaging array architecture.

line to the substrate, column line to row select lines, and output diffusion to substrate.

To read out a charge packet, the feedback capacitor C_{fb} is reset. The row select line is turned on and off, dumping the pixel charge onto the line. Since the column line capacitance is much greater than the pixel capacitance, nearly all of the imaged charge shifts onto the column line. This operation also resets the pixel. The pixel charge moves onto C_{fb} via the feedback action of the op-amp, producing a nominal output voltage of

$$V_{out} = V_{reset} - \frac{Q_{int}}{C_{fb}}. \tag{2.13}$$

Quantum Efficiency

The charge readout pixel has only one transistor, two wires, and two contacts, resulting in a relatively high optical aperture. The wires and transistor can be minimum size due to the small currents they pass.

Collection efficiency depends largely on doping profile. An enhanced CMOS process with additional implants could achieve collection efficiencies comparable to a CCD pixel. In a digital CMOS process, the choices are limited. Available photodiodes include the n^+ – p-well diode, p^+ – n-well diode, and n-well – substrate diode. Diffusion – well photodiodes will have relatively poor blue response because degenerate doping increases the defect density near the silicon surface; carriers generated near the surface tend to recombine before reaching the potential well. Red response is also compromised due to the limited space charge layer depth. The n-well photodiode avoids both of those problems, and should therefore produce good collection efficiency [37].

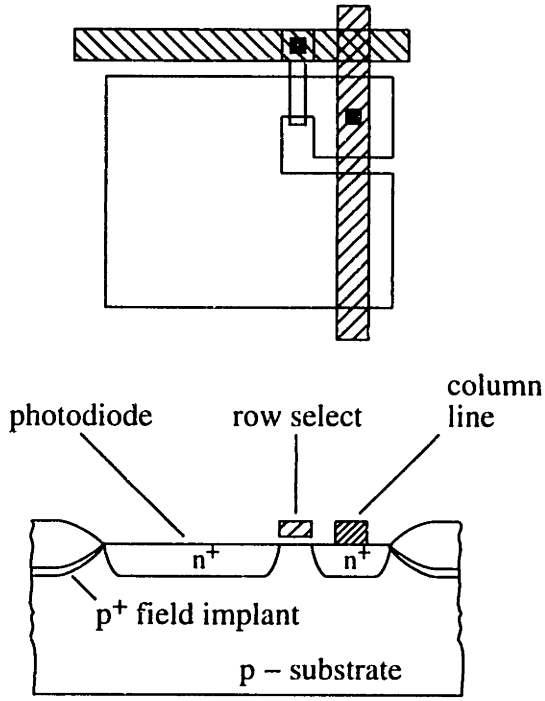


Figure 2-11: Charge readout pixel cross-section and plan view.

Conversion Gain

The conversion gain is C_{fb}^{-1}. There is no benefit in making the feedback capacitance greater than that which results in a full-scale swing for a saturation charge packet. Lower feedback capacitance results in higher conversion gain but limits the saturation charge.

Random Noise

Random noise sources include photonic noise, dark current shot noise, switch noise from the row select device, and amplifier noise. Photonic noise is a fundamental noise source for all imagers, and is caused by the Poisson statistics of the incoming photons. The resultant number of noise electrons is equal to the square root of the total number of imaged electrons. Dark current behaves the same as photogeneration current, and has the same shot noise characteristics. The main difference between these noise sources is that photonic noise increases with illumination, while dark current noise is independent of illumination.

Switch noise is the well known kTC noise caused by any switching operation onto a capacitor. The relevant capacitance in this case is the series combination of the pixel capacitance and the column line capacitance. Since the pixel capacitance is much less, the noise charge is very close

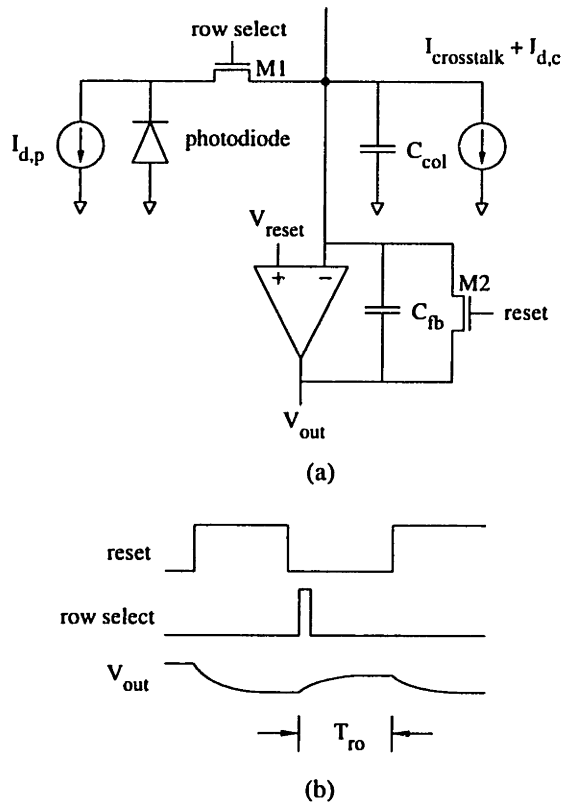


Figure 2-12: (a) Equivalent circuit for charge readout pixel and output circuit. $I_{crosstalk}$ is the optical crosstalk current from all pixels close to the column line. (b) Timing diagram for charge readout pixel. T_{ro} is the readout time.

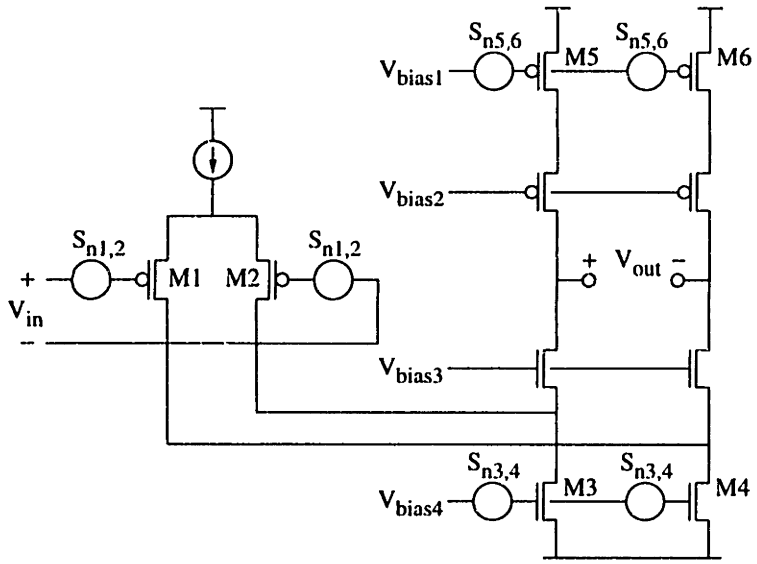


Figure 2-13: Random noise model for amplifier.

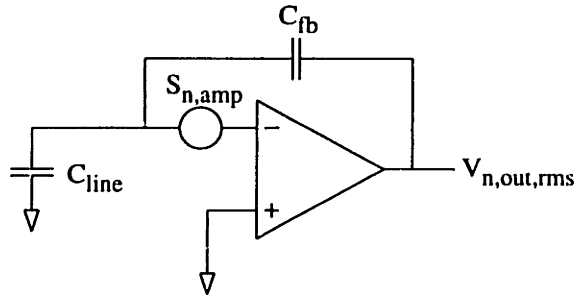


Figure 2-14: Random noise model for charge readout output circuit.

to

$$Q_{n,sw,rms} = \sqrt{kTC_{pixel}}. \tag{2.14}$$

It is desirable to keep the pixel capacitance as low as possible for noise reasons. Switch noise is independent of illumination. Note that the reset device M2 also produces switch noise, but it can be ignored because CDS mostly removes it.

The noise model for the amplifier (assumed to be a folded cascode design) is shown in Fig. 2-13. The input-referred noise spectral density is

$$S_n = 2\left(S_{n1,2} + S_{n3,4}\left(\frac{g_{m3,4}}{g_{m1,2}}\right)^2 + S_{n5,6}\left(\frac{g_{m5,6}}{g_{m1,2}}\right)^2\right). \tag{2.15}$$

The noise model for the entire readout circuit is shown in Fig. 2-14. The input-referred amplifier noise is multiplied by the closed loop gain of the system and limited by the system bandwidth. The noise, referred back to pixel charge, is

$$Q_{n,amp,rms} = C_{col}\sqrt{S_n \cdot \text{BW (Hz)} \cdot \pi/2}, \tag{2.16}$$

where BW is the closed-loop bandwidth of the system. This quantity is related to the readout time T_{ro}. The bandwidth should be just enough to allow settling in the readout time.

$$T_{ro} = N\tau = \frac{N}{2\pi \cdot \text{BW (Hz)}}, \tag{2.17}$$

where N settling times are assumed. For 10 bit settling, $N \approx 7$ is required to get within one lsb of the value at $t = \infty$. Amplifier noise decreases as T_{ro} increases. Readout time is usually limited by a minimum frame rate requirement.

The noise for a voltage readout pixel has a similar expression, where the noise source is the source follower device instead of the op-amp input devices. The noise is much lower for the voltage readout pixel because the corresponding capacitance is the pixel capacitance C_{pd} rather than the column line capacitance C_{col}. This noise is signal-independent.

Saturation Level

The pixel charge saturation level is the product of photodiode capacitance and photodiode voltage swing. Small pixel capacitance is desirable for low random noise, so high saturation level can only be achieved at the expense of higher random noise. The minimum photodiode voltage is approximately ground; if the voltage goes significantly below ground, the photodiode becomes forward biased. The maximum photodiode voltage is the reset voltage, which is probably about $V_{dd}/2$. Therefore, the voltage saturation level is no more than about $V_{dd}/2$, and will be less if the output buffer swing is the limiting factor.

Fixed-Pattern Noise

FPN in the charge readout pixel is caused by mismatches in the optical aperture A_{opt} and pixel dark current $I_{d,p}$. Sources outside the pixel include mismatches in column line capacitance C_{col}, op-amp offset voltage V_{os}, sense capacitance C_{fb}, op-amp gain A_{oa}, column dark current $I_{d,c}$, and M2's threshold voltage $V_{te,2}$ and M2's gate-drain overlap capacitance $C_{ol,s2}$. The photocurrent is I_p.

Table 2-2: FPN sources for a charge readout pixel

Parameter	Nominal Value	Sensitivity[1,2]	Mismatch
V_{os}	0 V	1	10 mV
A_{oa}	10,000	$C_{col}(K/A_{oa})^2(I_pT_{int}+q_0)$	10% A_{oa}
A_{opt}	50 μm^2	$(KT_{int}/A_{opt})I_p$	2% A_{opt}
$I_{p,d}$	0.5 fA	KT_{int}	10% $I_{p,d}$
C_{col}	2.5 pF	$-(K^2/A_{oa})(I_pT_{int}+q_0)$	0.1% C_{col}
$I_{d,c}$	20 fA	KT_{ro}	1% $I_{d,c}$
C_{fb}	14 fF	$-K^2(I_pT_{int}+q_0)$	1% C_{fb}
$V_{te,2}$	0.9 V	$KC_{ol,2}$	5 mV
$C_{ol,2}$	0.87 fF	$KV_{te,2}$	1% $C_{ol,2}$

[1]$K = A_{oa}/(A_{oa}C_{fb}+C_{col})$

[2]$q_0 = I_{d,p}T_{int}+I_{d,c}T_{ro}+C_{ol,2}V_{te,2}$

FPN can be calculated if the sensitivities of the output voltage to various mismatch sources are known. The relevant equation is

$$v_{FPN,rms}^2 = \left(\frac{\partial v_o}{\partial x_1}\right)^2 \sigma_{x1}^2 + \left(\frac{\partial v_o}{\partial x_2}\right)^2 \sigma_{x2}^2 + \cdots, \tag{2.18}$$

where x_1, x_2, $\cdots$ are the independent variables. The expression for output voltage is

$$V_{out} = V_{reset} + V_{os} + K(I_pT_{int}+q_0) \tag{2.19}$$

$$q_0 = I_{d,p}T_{int} + I_{d,c}T_{ro} + C_{ol,s2}V_{te,2} \tag{2.20}$$

$$K = \frac{A_{oa}}{A_{oa}C_{fb}+C_{col}}, \tag{2.21}$$

Table 2-2 shows the sensitivities, along with estimates of their expected error magnitudes, for a 10 μm $\times$ 10 μm pixel with 50% optical aperture. Mismatch magnitudes are for illustrative purposes only; mismatches (especially dark current) are likely to be highly process dependent.

The dominant sources are expected to be V_{os}, A_{opt}, $I_{d,p}$, and C_{fb} mismatch. Op-amp input offset voltage mismatch produces only an offset error, which can be removed by CDS. Optical aperture mismatch produces gain error. Pixel dark current mismatch produces an offset term which cannot be removed by CDS. Feedback capacitance mismatch produces offset and gain errors which cannot be removed by CDS. Figure 2-15 shows a graph of FPN for the charge readout pixel. Figure 2-16 is a graph of signal-to-noise ratio as a function of illumination.

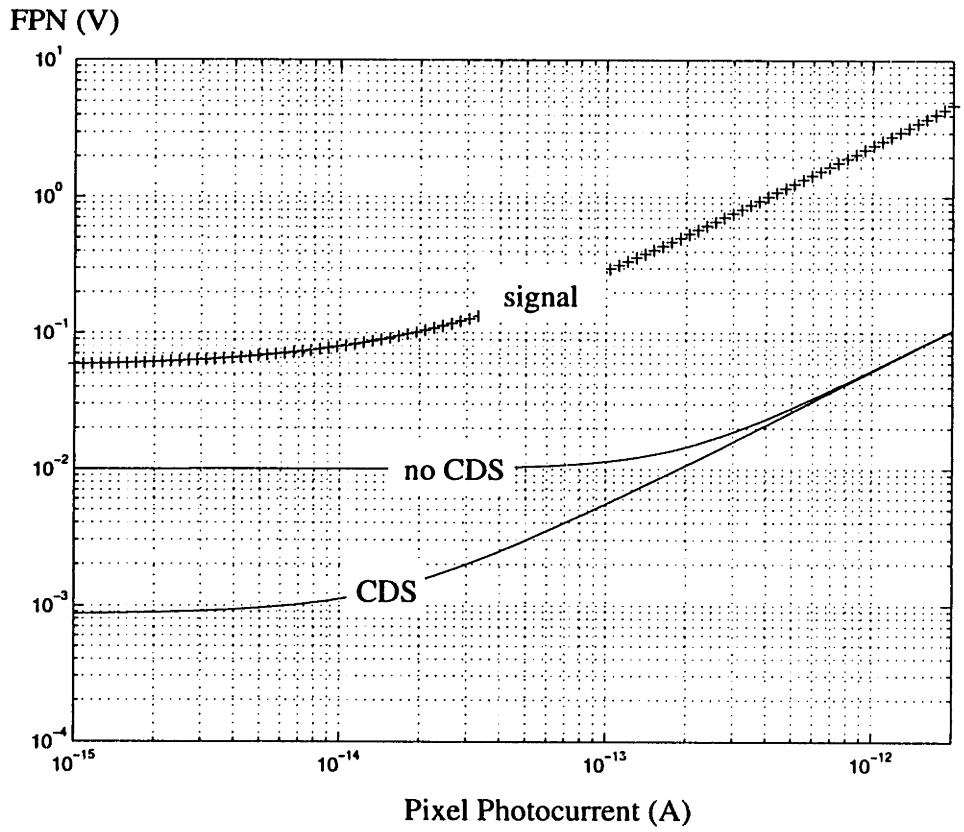


Figure 2-15: Simulated FPN for charge readout pixel with and without correlated double-sampling. The signal is plotted for reference.

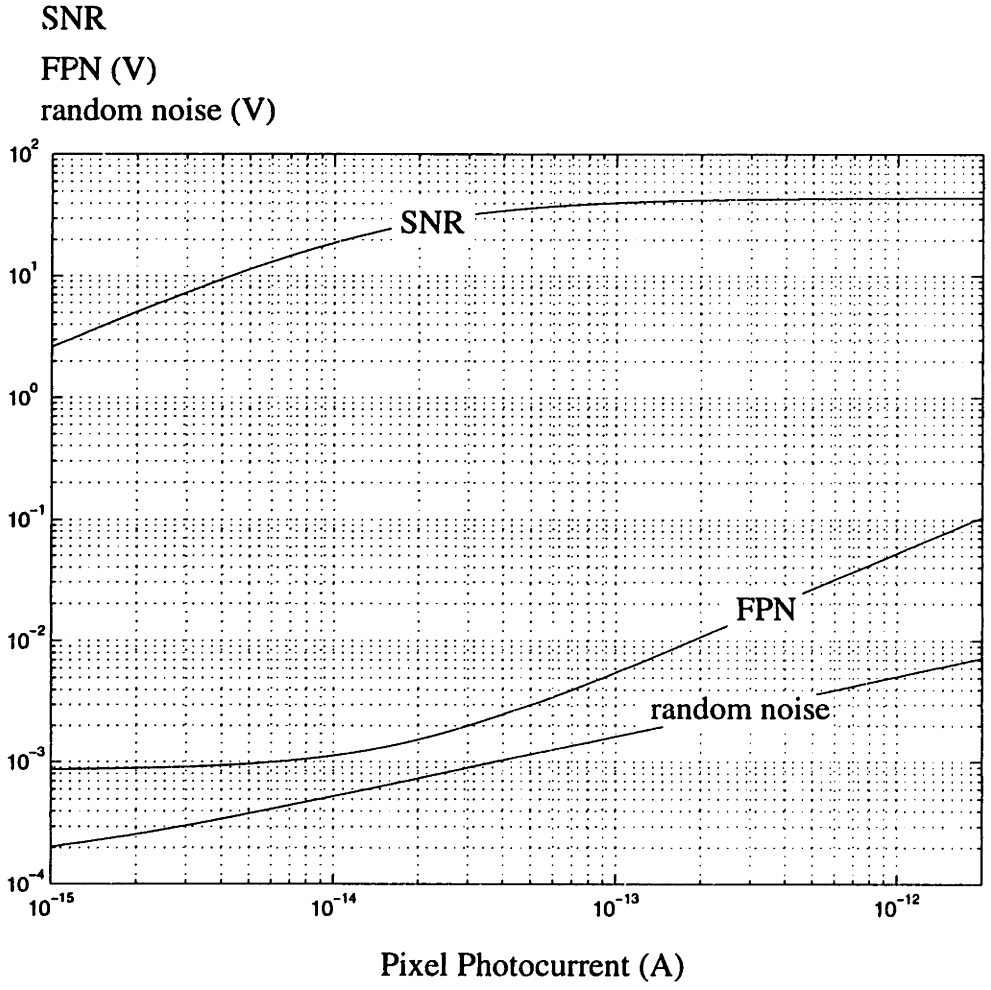


Figure 2-16: Simulated signal-to-noise ratio for charge readout pixel. Noise is sum of random and fixed-pattern noise.

Image Lag

Image lag depends on how the pixel is reset. If the row select is turned on just long enough to dump the charge onto the column line, and then turned off, the fraction of charge remaining on the pixel is

$$f = \frac{C_{pd}}{C_{pd} + C_{col}}. \tag{2.22}$$

If the row select remains off until the next readout time, the residual charge is seen as image lag. Image lag can be reduced by keeping the row select on until the op-amp has settled, moving essentially all imaged charge onto the feedback capacitor. It can also be reduced by using a separate reset phase. After reading out the charge, the row select remains on and the column line is connected to a voltage source.

Crosstalk

Electrical crosstalk is caused by incomplete settling and ground bounce. Incomplete settling may or may not be a problem, depending on how column lines are reset. The column line is at least partially reset by the amplifier moving pixel charge from the column line to the feedback capacitor. This process cannot be completed in finite time, so a residual charge is left on the column line. If the next pixel is read out immediately afterwards, its charge is combined with this residual charge, resulting in crosstalk. If the column line is reset with a voltage source before the next pixel is read out, there is no crosstalk.

Ground bounce is the variation in ground (substrate) potential due to current flow in the substrate. The difference in ground voltage between the pixel and the output circuit is indistinguishable from signal. Noise generators outside the pixel array can generally be shielded by surrounding substrate contacts; proper techniques and their effectiveness are discussed in several recent papers.

These techniques can usually not be directly applied to a pixel array because substrate contacts are seldom placed inside the array. At readout, the pixel charges dumped onto the column lines induce mirror charges to flow through the substrate and epitaxial layer. Figure 2-17 shows a cross-section through several column lines and photodiodes. The figure shows the substrate currents induced during readout due to column line #2. It can be seen that the effective ground for photodiode #1 depends on the currents induced by column line #2. The resultant crosstalk is extremely small since the total charge flowing in the substrate due to one pixel being read out is just the pixel charge, and the associated time constant is very small. Crosstalk is analyzed further for the voltage readout pixel.

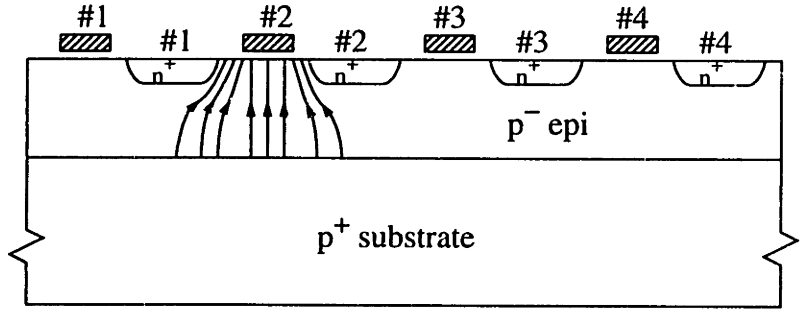


Figure 2-17: Substrate noise induced during readout.

Optical crosstalk is a more serious problem. Monochromatic light of wavelength λ entering an optical opening generates electron-hole pairs as an exponential function of depth into the silicon.

$$G(x) = \frac{I_\lambda}{x_0(\lambda)} e^{-x/x_0(\lambda)}, \tag{2.23}$$

where I_λ is the photon flux incident at the silicon interface ($x = 0$) and $x_0(\lambda)$ is the penetration depth at wavelength λ. For short wavelengths, the penetration depth is small. Light tends to get absorbed by surface layers such as polysilicon, and photons which do make it to the silicon frequently generate carriers so close to the surface that they are quickly recombined. Few carriers are generated beyond the photodiode space charge layer. For longer wavelengths, x_0 is large and many of the photons penetrate beyond the well to create electron-hole pairs in the neutral substrate. These charges diffuse randomly, some recombining, some entering the collection well, and some entering other potential wells. Charge entering adjacent pixels effectively results in a convolution of the actual image with a lowpass filter. The degradation of the image is minimal, and is common to all pixel arrays.

Unfortunately, optical crosstalk can also affect distant pixels in the same column in the charge readout design. The drain diffusion of the readout transistor is always connected to the column line. Charge must diffuse only a short distance before entering the drain, and it adds directly to the signal charge from the pixel currently being read out. This is a problem because the added charge comes from several pixels, which may be quite far from the readout pixel and have much higher illumination. This problem is particularly acute for an imager designed to operate over a very wide dynamic range. Optical crosstalk was a significant factor in the decision not to use the charge readout pixel.

Anti-blooming

Anti-blooming structures can be formed in several ways. A digital CMOS process can use a lateral overflow gate or p^+ – n-well photodiodes, using the n-well as a vertical overflow drain (VOD). If the n-well doping decreases with depth into the silicon, the n-well has a field directed towards the substrate. When a pixel saturates, it injects holes into the n-well, and the field directs them away from the surface. Because the field is probably low, the anti-blooming performance is poor. Extra implants can form a better VOD structure.

Another anti-blooming structure could be formed by using a p^+ S/D implant inside a separate n-well. By biasing the n-well at high potential, any stray holes in the substrate will be reflected away.

2.3.3 Current Output Pixel

A current output pixel transduces light to an output current. The output circuit converts the current to a voltage. There are many possible ways to do this, although many are impractical. For example, the direct pixel photocurrent could be switched into the column line during the readout time. The output circuit transresistance would have to be extremely high to get reasonable output signal levels, forming a slow time constant. It would be a difficult challenge to obtain even standard video rate operation. A poor phototransistor structure could be built using the n-well as base, the substrate as collector, and a p^+ S/D diffusion as emitter. This would increase the pixel output current level by a factor of β, but the n-well makes a poor base and the current gain is unlikely to exceed one order of magnitude.

Either of these two solutions tends to be noisy and/or slow to respond to the input light because of the continuous-time nature of the output. The output current can be sampled for only a short time (much less than the frame time). If the detector responds quickly, the sampled noise is averaged only over the sampling interval and is therefore relatively high. If the detector is slow, noise is effectively averaged over a long time, but the slow response produces image lag.

A better implementation is shown in Fig. 2-19 [38]. The photodiode is reset at the start of the integration period by turning on M2 and M3, and the switch to the column bias current is closed. The photodiode voltage rises to the level required to sink I_{bias} through M1. Next, Devices M2 and M3 are turned off and the current bias switch is opened. Charge integrates on the photodiode, dropping its potential V_{pd}. At the end of the integration period, the row select device M3 turns on and the switch to the transresistance circuit is closed. The output current, I_{out}, has a quadratic dependence on the integrated charge. The transresistance element converts

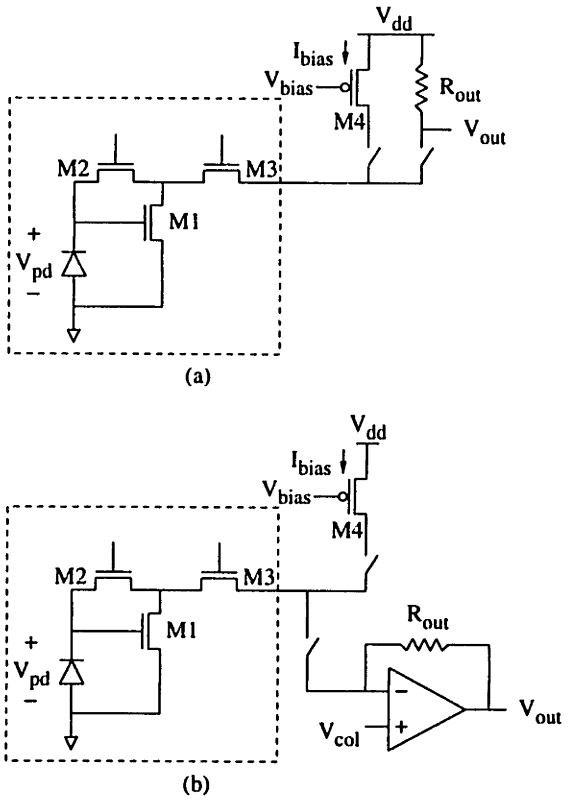


Figure 2-18: Alternative output circuits for current output pixel.

the output current I_{out} to an output voltage V_{out}. In this case, the diode-connected transistor M5 performs this role. A differential readout mode is possible where the direct current I_{out} and the bias current I_{bias} are read out simultaneously, permitting a correction for column-to-column variation in the bias current.

Two alternative implementations of the output circuit are shown in Fig. 2-18. In Fig. 2-18(a), the output current passes through a resistor, and the voltage produced is proportional to current. In Fig. 2-18(b), the op-amp holds the column line at a constant voltage, and forces the pixel output through the feedback resistor. With either of these output circuits, the response to illumination is nonlinear, since the relationship between illumination and current is nonlinear, and voltage is proportional to current.

Quantum Efficiency

The current readout pixel in Fig. 2-19 requires three transistors, four wires, and five contacts, so the optical aperture for this pixel is likely to less than for a similar size charge readout

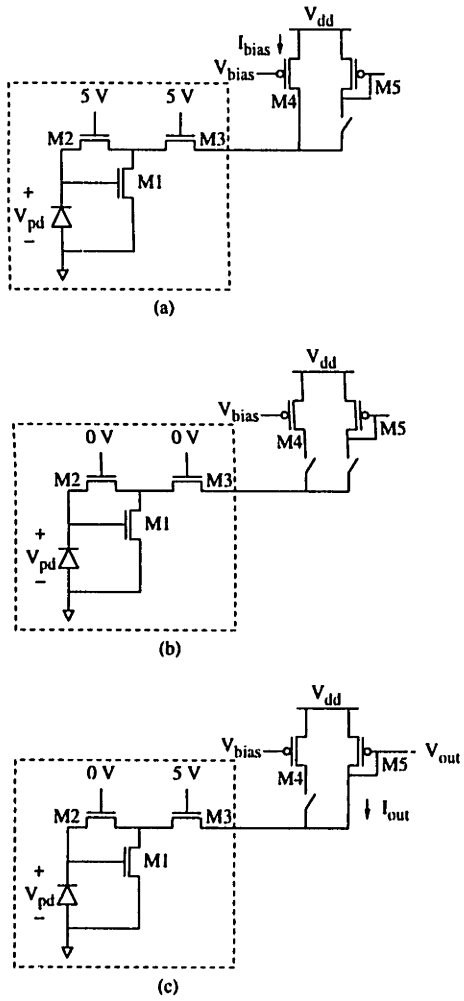


Figure 2-19: Current output pixel and operating sequence. Portion in dashed box is inside the pixel, while devices M4 and M5 are common to all devices in the column. (a) Reset phase. (b) Charge integration phase. (c) Readout phase.

pixel. Collection efficiency will also tend to be less due to the presence of extraneous n^+ S/D diffusions such as M1's source and drain. The photodiode choices are the same as for a charge readout pixel. The S/D diffusion in a digital CMOS process provides fair collection efficiency; an enhanced CMOS process allows specialized doping profiles for increased collection efficiency.

Conversion Gain

Conversion occurs in three steps. The integrated charge generates a voltage on the sense capacitance, this voltage generates an output drain current through M1, and M1's drain current is converted to voltage by the transresistance of the output circuit. The conversion gain changes over the integration time; the conversion gain as a function of photodiode voltage is

$$\frac{1}{C_{pd}} \cdot k_n \frac{W_1}{L_1}(V_{pd} - V_T)^2 \cdot R_{out}, \tag{2.24}$$

where V_{pd} is voltage across the photodiode, R_{out} is the transresistance of the output circuit, and C_{pd} is the capacitance of the photodiode. Conversion gain can be selected by the choice of W_1, L_1, and R_{out}.

Random Noise

Random noise is the sum of the noises introduced at reset, during charge integration, and at readout. Noise sources at reset include photodiode reset noise, thermal noise in M1 and M4, and $1/f$ noise in M1 and M4. Photodiode reset noise has been discussed previously; it produces a mean square voltage at M1's gate of

$$v_{n,rms}^2 = \frac{kT}{C_{pd}}. \tag{2.25}$$

It will be assumed that $1/f$ noise is sufficiently correlated between the reset and output phases that correlated double-sampling removes it. The remaining noise sources are thermal noise in M1 and M4. A noise model for the reset phase is shown in Fig. 2-20(a). Thermal noise referred to M1's gate has mean square voltage

$$v_{n,rms}^2 = \frac{8}{3}\frac{kT}{g_{m1,rst}} \cdot \text{BW (Hz)} + \frac{8}{3}\frac{kT}{g_{m4,rst}} \cdot \text{BW (Hz)} \left(\frac{g_{m4,rst}}{g_{m1,rst}}\right)^2 \tag{2.26}$$

$$= \frac{8}{3}\frac{kT}{g_{m1,rst}}\left(1 + \frac{g_{m4,rst}}{g_{m1,rst}}\right) \cdot \frac{1}{2\pi}\frac{g_{m1,rst}}{C_{col}}\frac{\pi}{2} \tag{2.27}$$

$$= \frac{2}{3}\frac{kT}{C_{col}}\left(1 + \frac{g_{m4,rst}}{g_{m1,rst}}\right) \tag{2.28}$$

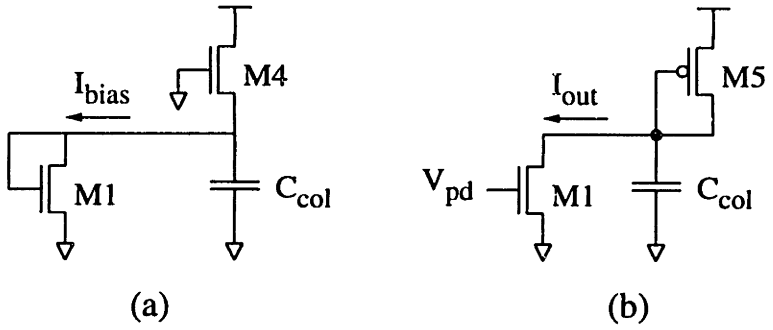


Figure 2-20: Noise models for current readout pixel. (a) Reset phase. (b) Sample phase.

The only noise source during the charge integration period is dark current shot noise, which has been previously discussed. The mean square noise referred to M1's gate is

$$v_{n,rms}^2 = \frac{T_{int} I_d q_e}{C_{pd}^2}, \tag{2.29}$$

where q_e is the charge of one electron, or 1.6×10^{-19}C.

The noise sources in the readout phase are thermal and $1/f$ noise in M1 and M5. As before, $1/f$ noise is assumed to be removed using correlated double-sampling. The noise model for the pixel and readout circuit in the readout phase is shown in Fig. 2-20(b). The voltage noise referred to M1's gate is

$$v_{n,rms}^2 = \frac{8}{3}\frac{kT}{g_{m1,out}} \cdot \text{BW (Hz)} + \frac{8}{3}\frac{kT}{g_{m5,out}} \cdot \text{BW (Hz)} \left(\frac{g_{m5,out}}{g_{m1,out}}\right)^2 \tag{2.30}$$

$$= \frac{8}{3}\frac{kT}{g_{m1,out}} \left(1 + \frac{g_{m5,out}}{g_{m1,out}}\right) \cdot \frac{1}{2\pi} \frac{g_{m5,out}}{C_{col}} \frac{\pi}{2} \tag{2.31}$$

$$= \frac{2}{3}\frac{kT}{C_{col}} \left(\frac{g_{m5,out}}{g_{m1,out}} + \left(\frac{g_{m5,out}}{g_{m1,out}}\right)^2\right) \tag{2.32}$$

The total square rms noise is the sum of the individual contributions.

$$q_{n,rms}^2 = T_{int} I_d q_e + kTC_{pd} + \frac{2}{3} kT \frac{C_{pd}^2}{C_{col}} \left(1 + \frac{g_{m4,rst}}{g_{m1,rst}} + \frac{g_{m5,out}}{g_{m1,out}} + \left(\frac{g_{m5,out}}{g_{m1,out}}\right)^2\right) \tag{2.33}$$

$$\approx T_{int} I_d q_e + kTC_{pd} \tag{2.34}$$

where the approximation follows because $C_{col} \gg C_{pd}$.

Saturation Level

The saturation level expressed in terms of charge is the product of the sense capacitance and photodiode voltage swing. The lowest usable photodiode voltage is nominally M1's threshold voltage. As the voltage decreases further and M1 enters subthreshold, the output current decreases so much that the column line cannot settle in the time allotted for the readout phase. A further problem in the case of the readout circuits of Fig. 2-18 is that the subthreshold current is so low that the output voltage becomes indistinguishable from V_{dd}. The maximum photodiode voltage is

$$V_{pd,max} = V_{T1} + \sqrt{\frac{2I_{bias}}{k_n(W/L)}}, \tag{2.35}$$

and occurs at reset. The voltage swing is therefore

$$\Delta V_{pd} = \sqrt{\frac{2I_{bias}}{k_n(W/L)}}. \tag{2.36}$$

FPN

FPN sources include mismatches in bias current, output circuit transresistance R_{out}, M1's width and length W_1 and L_1, M2's overlap capacitance $C_{ol,2}$ and effective threshold voltage $V_{te,2}$, photodiode capacitance C_{pd}, and optical aperture A_{opt}. Threshold voltage mismatch in M1 is cancelled, to first order, by the reset procedure.

Bias Current A mismatch ΔI_{bias} in bias current causes a mismatch Δq in photodiode charge of magnitude

$$\Delta q = \left(C_{pd} + C_{ol,2} + \frac{2}{3}W_1L_1C_{ox}\right)\frac{\Delta I_{bias}}{g_{m1,rst}} \tag{2.37}$$

$$= C_{sense}\frac{I_{bias}}{g_{m1,rst}}\frac{\Delta I_{bias}}{I_{bias}}. \tag{2.38}$$

during the reset phase. This produces offset FPN.

Output Circuit Transresistance A fractional mismatch δ in transresistance R_{out} is equivalent to a fractional mismatch δ in pixel output current, which can be referred back to a photodiode voltage mismatch by dividing by g_{m1}.

$$\Delta q = C_{sense}\,\Delta V_{pd} \tag{2.39}$$

$$= C_{sense}\frac{\Delta I_{out}}{g_{m1,out}} \tag{2.40}$$

$$= C_{sense} \frac{\Delta R_{out}}{R_{out}} \frac{I_{out}}{g_{m1,out}} \tag{2.41}$$

M1 Aspect Ratio A fractional decrease δ in M1's W/L ratio is equivalent to a fractional decrease δ in I_{bias} combined with a fractional increase δ in R_{out}. In the reset phase, the photodiode voltage V_{pd} would be the same, because M1 would be operating at the same current density. In the readout phase, the output current would be lower by a fraction δ, but would produce the same output voltage because the transresistance is higher by a fraction δ. Combining these using Equations 2.38 and 2.41 gives

$$\Delta q = C_{sense} \frac{\Delta(W_1/L_1)}{W_1/L_1} \left(\frac{I_{out}}{g_{m1,out}} - \frac{I_{bias}}{g_{m1,rst}} \right) \tag{2.42}$$

$$= C_{sense} \left(\frac{I_{out}}{g_{m1,out}} - \frac{I_{bias}}{g_{m1,rst}} \right) \left(\frac{\Delta W_1}{W_1} - \frac{\Delta L_1}{L_1} \right) \tag{2.43}$$

M1 aspect ratio mismatch produces only a gain-like FPN, *i.e.*, FPN which is approximately proportional to illumination, as opposed to offset FPN which has the same magnitude at every illumination.

M2 Threshold Voltage and Overlap Capacitance Mismatches in M2's effective threshold voltage and overlap capacitance directly inject charge onto the photodiode, producing an offset FPN of magnitude

$$\Delta q = -V_{te,2} \Delta C_{ol,2} - C_{ol,2} \Delta V_{te,2}, \tag{2.44}$$

where $V_{te,2}$ is the effective threshold voltage of M2, including back-gate effect, $\Delta V_{te,2}$ is the rms mismatch in M2's threshold voltage, $C_{ol,2}$ is M2's gate-drain overlap capacitance, and $\Delta C_{ol,2}$ is the rms mismatch in M2's gate-drain overlap capacitance.

Charge Sense Capacitance A fractional increase δ in charge sense capacitance is equivalent to a fractional decrease δ in the integrated charge, leading to a gain FPN

$$\Delta q = -q \frac{\Delta C_{sense}}{C_{sense}} \tag{2.45}$$

$$= -\frac{q}{C_{sense}} \left(\Delta C_{pd} + \Delta C_{ol,2} + \frac{2}{3} W_1 L_1 C_{ox} \left(\frac{\Delta W_1}{W_1} + \frac{\Delta L_1}{L_1} \right) \right), \tag{2.46}$$

where C_{sense} is the total charge sense capacitance, ΔC_{sense} is the rms mismatch in sense capacitance, C_{pd} is the photodiode capacitance, and ΔC_{pd} is the rms mismatch in photodiode capacitance.

Table 2-3: FPN sources for a current readout pixel.

Parameter	Value	Sensitivity ($\Delta Q/\Delta x$)	Mismatch
I_{bias}	120 μA	$\frac{C_{sense}}{g_{m1,rst}}$	1% I_{bias}
R_{out}	33 KΩ	$\frac{C_{sense} I_{out}}{g_{m1,out} R_{out}}$	1% R_{out}
W_1	2 μm	$-q\frac{2}{3}\frac{L_1 C_{ox}}{C_{sense}} + \frac{C_{sense}}{W_1}\left(\frac{I_{out}}{g_{m1,out}} - \frac{I_{bias}}{g_{m1,rst}}\right)$	0.1 μm
L_1	0.8 μm	$-q\frac{2}{3}\frac{W_1 C_{ox}}{C_{sense}} - \frac{C_{sense}}{L_1}\left(\frac{I_{out}}{g_{m1,out}} - \frac{I_{bias}}{g_{m1,rst}}\right)$	0.1 μm
$V_{te,2}$	0.9 V	$-C_{ol,2}$	5 mV
$C_{ol,2}$	0.87 fF	$-V_{te,2} - \frac{q}{C_{sense}}$	2% $C_{ol,2}$
C_{pd}	10 fF	$-\frac{q}{C_{sense}}$	2% C_{pd}
A_{opt}	30 μm^2	$\frac{q}{A_{opt}}$	2% A_{opt}
I_d	0.5 fA	T_{int}	10% I_d

Optical Aperture Quantum efficiency, and therefore integrated charge, is proportional to optical aperture. Optical aperture is defined as the fraction of the pixel area which is optically active. A fractional increase δ in optical aperture produces the same effect as a fractional increase δ in illumination. Given a constant integration period, the integrated charge is increased by a fraction δ.

$$\Delta q = q\frac{\Delta A_{opt}}{A_{opt}}, \tag{2.47}$$

where A_{opt} is the optical aperture and ΔA_{opt} is the rms mismatch in optical aperture.

Optical aperture may be correlated with charge sense capacitance, but this is not necessarily so. It is possible to use a field oxide region uncovered by polysilicon or metal to generate photocharge, which is then collected by a small S/D diffusion. The advantages are small sense capacitance and improved responsivity. The optical aperture in this case is not defined by the photodiode area, and thus their mismatches are uncorrelated. Even if they are correlated, the correlation could be positive or negative, depending on pixel construction and process details.

A graph of FPN as a function of pixel photocurrent is shown in Fig. 2-21. Mismatch parameters are from Table 2-3. A simple square-law model for transistor current is assumed, using a t_{ox} of 17.2 nm and a μ_n of 500 cm^2/V·s. Also, the output transresistance is assumed to be a simple resistor. With CDS, M2's overlap capacitance is the dominant source of FPN at low illumination. Note that without CDS, FPN approaches zero as signal approaches saturation.

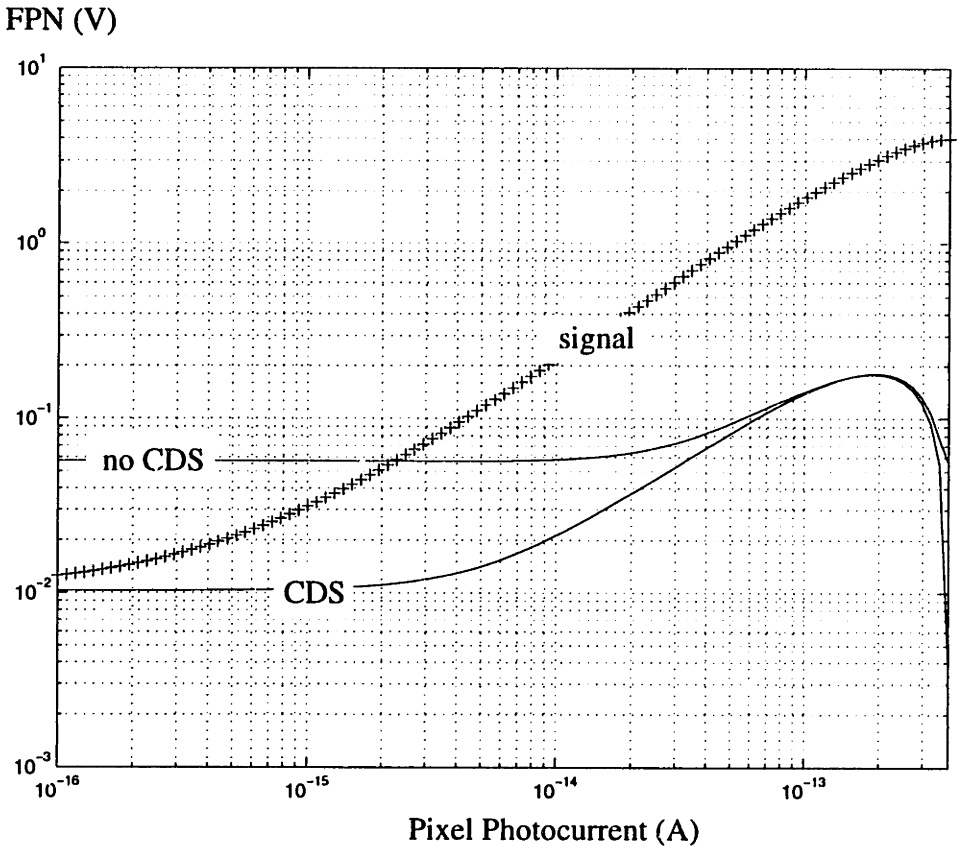


Figure 2-21: Simulated FPN and signal for current output pixel.

For high illumination, the photodiode voltage V_{pd} is close to ground; M1 passes very little current and has very low transconductance. Stated another way, whether M1's gate-source voltage is 0.1 V or 0.11 V, the output voltage is very close to V_{dd}. Signal-to-noise ratio is plotted in Fig. 2-22.

Image Lag

During pixel reset, the column line voltage must settle from its value at readout to the steady-state value which, when applied to the gate of M1, produces a drain current of I_{bias}. Since the time allowed for the system to settle must be limited, the system does not completely settle. The residual voltage on the column line at the end of the reset phase is trapped on the photodiode, and affects that pixel's output current on the next frame.

The magnitude of trapped voltage depends on the type of output circuit that is used. For the output circuits of Figs. 2-19 and 2-18(a), the voltage on the column line during readout

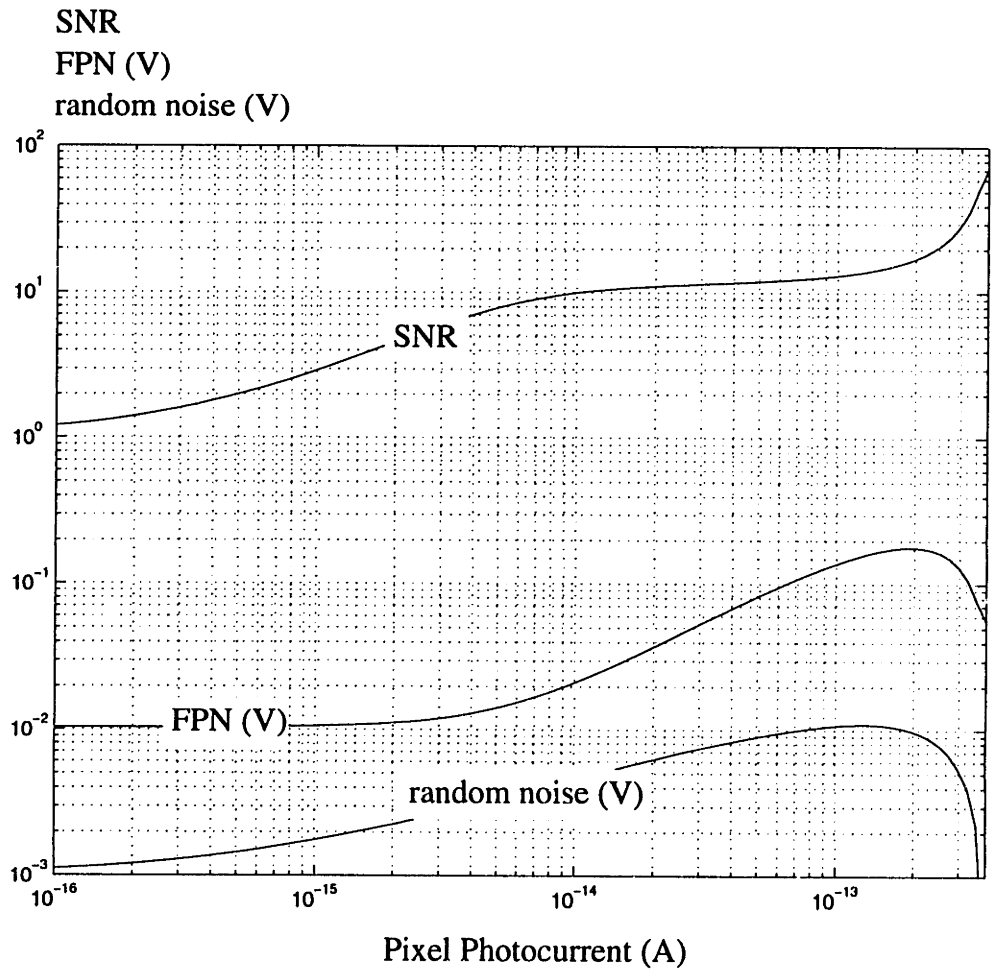


Figure 2-22: Simulated signal-to-noise ratio for current readout pixel. Noise is sum of random and fixed-pattern noise.

varies considerably with output current level. For example, with an I_{bias} of 1 μA and an R_{out} of 1 MΩ, the circuit of Fig. 2-18(a) produces a 1 V full-scale swing. The residual voltage at the end of the reset phase is

$$v_{residual} = (v_{out} - v_{rst})e^{-t_r/\tau_r}, \tag{2.48}$$

where v_{out} is the output voltage during the readout phase, v_{rst} is the column line voltage during reset in the infinite time limit, t_r is the reset phase time, and τ_r is the time constant during the reset phase, given by

$$\tau_r = \frac{C_{col}}{g_{m1}}. \tag{2.49}$$

For the readout circuit of Fig. 2-18(b), the op-amp holds the column line voltage close to V_{col}, and therefore the column line voltage during the readout phase is not dependent on output current level. During the reset phase, the voltage on the column line is

$$v_{col} = (V_{col} - v_{rst})e^{-t_r/\tau_r}, \tag{2.50}$$

which does not depend on the output current. The residual voltage trapped on the photodiode is therefore not dependent on output current level, and does not appear as lag.

Crosstalk

Electrical crosstalk caused by substrate current will be analyzed for the voltage readout pixel. The result there, which extends to the current readout pixel, is that crosstalk between adjacent lines is negligible. This pixel has no other obvious source of electrical crosstalk.

Optical crosstalk occurs when light entering a pixel generates photocharge outside the pixel's collection region. Optical crosstalk occurs between pixels and from pixels onto the column line. The level of the crosstalk current will usually be much smaller than the signal currents; the bias current in the actual devices was around 120 μA. The only possible exception is for very bright pixels, where the output signal current is close to zero. Crosstalk current will cause those pixels appear less bright than they actually are.

Anti-blooming

A charge output pixel can use any of the anti-blooming options available to the charge readout pixel. The charge readout pixel inherently has two forms of anti-blooming. The reset transistor also acts as a lateral overflow gate if its "off" voltage is not too low. The column lines also provide some anti-blooming, since the S/D diffusions attached to column lines act as sinks for

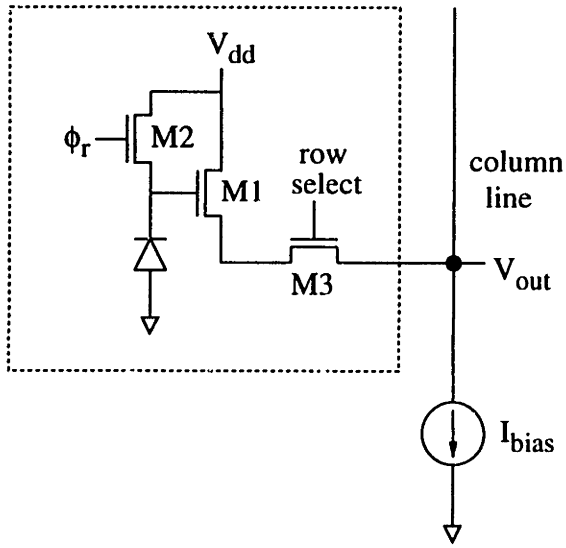


Figure 2-23: Voltage output pixel.

blooming charge. The pixel blooming current is much smaller than the pixel output current, so the blooming current does not result in significant crosstalk.

2.3.4 Voltage Output Pixel

The basic form of the voltage readout pixel is shown in Fig. 2-23. The pixel is reset by pulling ϕ_r high. In a "hard" reset, ϕ_r is pulled more then a threshold drop above V_{dd} and the photodiode is reset to V_{dd}. In a "soft" reset, ϕ_r does not go high enough to equalize the voltage across M2. Assuming ϕ_r goes to V_{dd}, the photodiode is reset to approximately $V_{dd} - V_{te,2}$. Integration starts when ϕ_r goes low. The photodiode potential drops at a rate proportional to the illumination. At the end of the integration period, the row select device M3 turns on, connecting M1's source to the current source at the bottom of the column line. The output signal is the voltage on the column line, which has an approximate linear dependence on the illumination.

Quantum Efficiency

The optical aperture of a voltage readout pixel should be similar to that of a current readout pixel. Like that pixel, the voltage readout pixel contains three transistors, at least four wires (column line, row select, V_{dd}, and ϕ_r), and five contacts. Typical optical apertures are around 20-30% [26].

Similar remarks apply to the collection efficiency. In a digital process, only S/D diffusion

and n-well photodiodes are available. An enhanced CMOS process can have a customized photodiode doping profile for improved collection efficiency.

Conversion Gain

The conversion gain is $G_{sf}C_{sense}^{-1}$, where G_{sf} is the source follower gain and the charge sense capacitance C_{sense} is the total capacitance at the gate of M2. Since C_{sense} is approximately proportional to the photodiode area, the output voltage for a given light intensity is fixed by process parameters. It may be desirable to increase the in-pixel gain for improved desensitivity to external noise and offset. One common technique is to have separate charge integration and charge sense nodes. If the photodiode and charge sense areas are A_{pd} and A_{cs}, respectively, the conversion gain is increased by a factor of A_{pd}/A_{cs}. In a photodiode implementation, this results in lag since the current from the photodiode into the charge sense node decreases gradually after the light is removed. A photogate implementation does not have the same problem.

Random Noise

Random noise sources include switch noise, the photonic and dark current shot noise, and thermal and $1/f$ noise in the source follower transistor. These are caused by independent phenomena, and therefore they can be added in quadrature. Switch noise depends on whether the reset is hard or soft; a hard reset will be assumed here, and the soft reset case will be discussed when the wide dynamic range pixel is analyzed. Using an analysis similar to that of the current output pixel, the total random noise is

$$Q_{n,rms}^2 = kTC_{sense} + q(I_d + I_p)T_{int} + \frac{8}{3}C_{sense}^2 \frac{kT}{g_{m1}} \frac{g_{m1} + g_{mb1}}{2\pi C_{col}} \frac{\pi}{2} \tag{2.51}$$

$$= kTC_{sense} + q(I_d + I_p)T_{int} + \frac{2}{3}\left(1 + \frac{g_{mb1}}{g_{m1}}\right)\frac{C_{sense}}{C_{col}}\, kTC_{sense} \tag{2.52}$$

$$\approx kTC_{sense} + q(I_d + I_p)T_{int}, \tag{2.53}$$

where T_{int} is the integration period, I_d is the dark current, and I_p is the photogenerated current. The random noise, to a close approximation, is the sum of reset noise and dark + photonic shot noise.

Saturation Level

The voltage saturation level of the photodiode is the difference between its maximum and minimum levels. The maximum photodiode voltage V_{reset} is achieved at reset, and will typically

be around 4 V. The lower limit on the photodiode voltage is set by the current source compliance; the photodiode voltage must be above the minimum operating voltage of the current source by at least a threshold drop plus the overdrive of M1. This is usually around 1.5 V. The saturation level is at most 2.5 V.

In terms of illumination, the saturation level is proportional to the voltage saturation level and the charge sense capacitance, and inversely proportional to the quantum efficiency, optical aperture, and integration time. For example, the dynamic range extension technique described in this work increases the saturation level by using shorter integration times for higher illuminations.

Fixed-Pattern Noise

FPN sources include mismatches in photodiode capacitance C_{pd}, optical aperture A_{opt}, bias current I_{bias}, pixel dark current I_d, M1's width and length W_1 and L_1, M1's threshold voltage $V_{te,1}$, M2's overlap capacitance $C_{ol,2}$, and M2's threshold voltage $V_{te,2}$. With the reset device M2 on, the output voltage is $V_{out,rst}$. At the end of the integration period, clock feedthrough from the reset transistor, photocurrent, and dark current have dropped the voltage on the sense capacitance. The voltage at the output is

$$\textbf{hard reset: } V_{out} = V_{out,rst} - \frac{G_{sf}}{C_{sense}}(C_{ol,2}(V_{dd} + V_{te,2}) + (I_p + I_d)T_{int}) \tag{2.54}$$

$$\textbf{soft reset: } V_{out} = V_{out,rst} - \frac{G_{sf}}{C_{sense}}(C_{ol,2}V_{dd} + (I_p + I_d)T_{int}) \tag{2.55}$$

where the sense capacitance is

$$C_{sense} = C_{pd} + C_{ol,2} + \frac{2}{3}C_{ox}W_1L_1(1 - G_{sf}), \tag{2.56}$$

the source follower gain is

$$G_{sf} = \frac{g_{m1}}{g_{m1} + g_{mb1} + g_{o1}}. \tag{2.57}$$

The transconductance of M1 is g_{m1}, the back-gate transconductance of M1 is g_{mb1}, and the drain-source conductance of M1 is g_{o1}. Table 2-4 shows the FPN sources, their output voltage sensitivities, and their expected mismatches assuming a hard reset is used.

The dominant sources of gain FPN are expected to be C_{pd} and A_{opt}. The dominant sources of offset FPN (without CDS) are $V_{te,1}$ and $C_{ol,2}$. The only source of offset FPN with CDS is the dark current. Current source mismatch is the only source of column FPN, but fortunately V_{out} is fairly insensitive to I_{bias}, and the offset is easily removed by CDS. Estimated rms FPN

Table 2-4: FPN sources for a voltage readout pixel.

Parameter	Value	Sensitivity	Mismatch
C_{pd}	10 fF	$\frac{G_{sf}(I_p T_{int}+q_0)}{C_{sense}^2}$	2% C_{pd}
A_{opt}	30 μm^2	$-\frac{G_{sf} I_p T_{int}}{C_{sense} A_{opt}}$	2% A_{opt}
I_{bias}	1 μA	$-(g_{mb1} + g_{m1} c_{ratio})^{-1}$	1% I_{bias}
I_d	0.3 fA	$-\frac{G_{sf} T_{int}}{C_{sense}}$	10% I_d
W_1	2 μm	$\frac{2}{3}\frac{L_1 C_{ox} G_{sf}(1-G_{sf})(I_p T_{int}+q_0)}{C_{sense}^2} + \frac{I_{bias}}{W_1(g_{mb1}+g_{m1} c_{ratio})}$	0.1 μm
L_1	0.8 μm	$\frac{2}{3}\frac{W_1 C_{ox} G_{sf}(1-G_{sf})(I_p T_{int}+q_0)}{C_{sense}^2} - \frac{I_{bias}}{L_1(g_{mb1}+g_{m1} c_{ratio})}$	0.1 μm
$V_{te,1}$	1.2 V	-1	5 mV
$C_{ol,2}$	0.87 fF	$\frac{G_{sf}(I_p T_{int}+q_0)}{C_{sense}^2} - \frac{G_{sf}(V_{dd}+V_{te,2})}{C_{sense}}$	2% $C_{ol,2}$
$V_{te,2}$	1.0 V	$-G_{sf}\frac{C_{ol,2}}{C_{sense}}$	5 mV

$q_0 = I_{d,p} T_{int} + C_{ol,2}(V_{dd} + V_{te,2})$ $\qquad$ $c_{ratio} = \frac{C_{ol,2}+C_{pd}}{C_{ol,2}+C_{pd}+C_{gs1}}$

is plotted as a function of pixel photocurrent in Fig. 2-24. The signal-to-noise ratio is plotted in Fig. 2-25.

The major difference between the current output and voltage output FPN plots is their behavior at high illumination. The current output pixel FPN becomes very small at high illumination because the output transistor M1 is cutoff. When this happens, no current flows and the output voltage is simply V_{dd}, regardless of any device parameters. The voltage output pixel does not experience any dramatic change at high illumination.

Another difference is that the voltage output pixel appears to have much better FPN performance. The main reason is the manner in which correlated-double sampling is assumed to operate. In the current output pixel, correlated-double sampling removes offsets due to bias current and output resistance, but not charge feedthrough from the reset switch. For the voltage output pixel, charge feedthrough can be removed by sampling the reset voltage immediately after the pixel has been reset and the reset gate turned off. The residue offset FPN is due solely to dark current mismatch.

The voltage output pixel should have lower gain FPN because it uses fewer signal conversions. The voltage output pixel converts light to charge, charge to voltage, and voltage to voltage. The current output pixel converts light to charge, charge to voltage, voltage to cur-

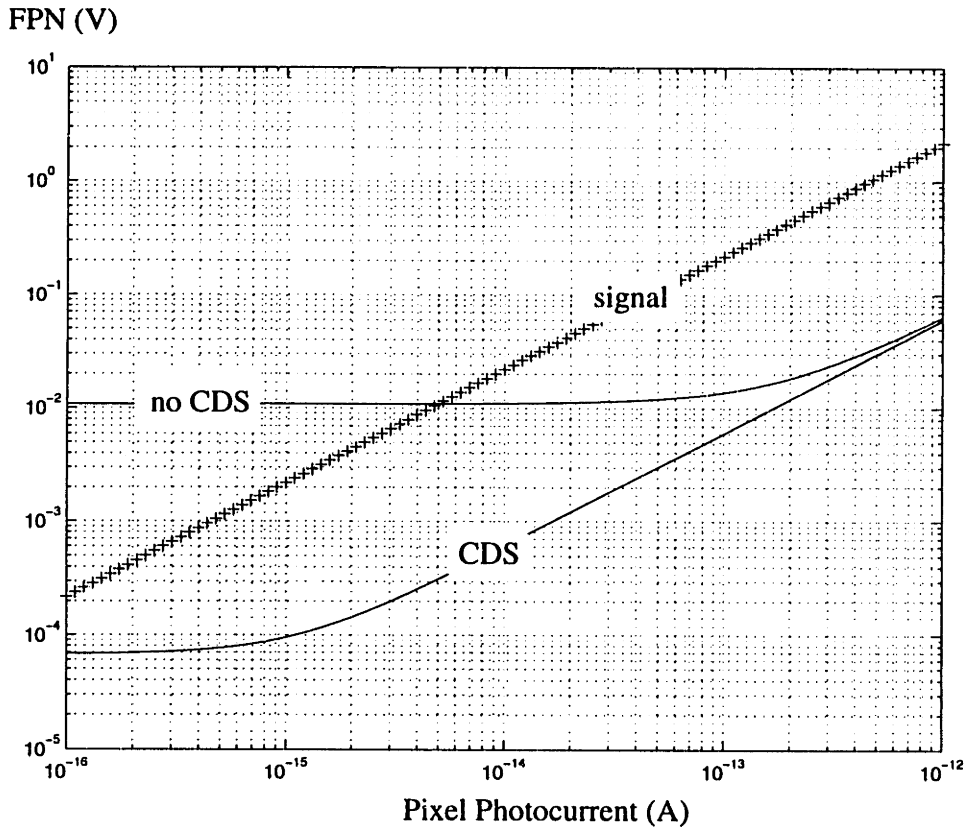


Figure 2-24: Simulated FPN and signal for voltage readout pixel.

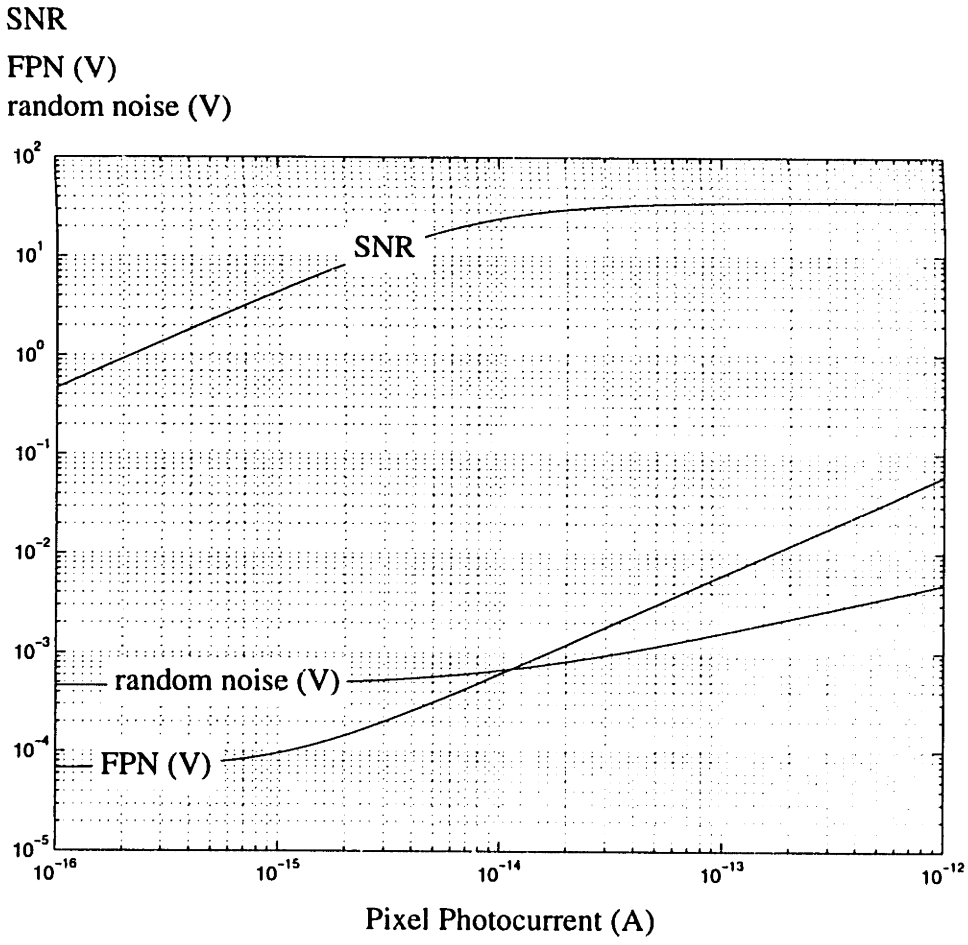


Figure 2-25: Simulated SNR for voltage readout pixel. Noise is the sum of random and fixed-pattern noise (with CDS).

rent, and current to voltage. Each of these conversions introduces a gain term which can vary between pixels.

Image Lag

Image lag is caused by the incomplete reset of the photodiode during a soft reset. As the photodiode voltage approaches its asymptotic reset value, M2's current decreases and the settling time constant increases. Since there is a limited time available for reset, the photodiode voltage may not completely settle. The dependence of the residual voltage on the previous illumination level is seen as image lag. A pixel which uses a hard reset should not have significant image lag because the settling time constant has a lower limit.

Crosstalk

The voltage readout pixel has many of the same crosstalk sources as the charge readout pixel. Optical crosstalk has a minimal effect on the output voltage. Although charge can still enter the column line from pixels which are not being read out, it has only a very small effect on the total current in the source follower device, and a negligible effect on its output voltage. For example, 1 pA would be an enormous amount of crosstalk current from one pixel into the column line. If all pixels in a column have that amount of crosstalk, the total crosstalk current is still less than 1 nA, which is insignificant compared to the a typical 1 μA bias current. The voltage readout pixel array is still susceptable to pixel-to-pixel optical crosstalk, but as stated earlier this is relatively benign, and unavoidable in any case.

Incomplete settling on the column line can cause electrical crosstalk. When one sample is being read out, the column line voltage has a component due to the previous sample. This residual component decays quite fast with reasonable source follower bias currents, and can be completely eliminated by shorting the column line to a fixed voltage source between samples.

Electrical crosstalk through the substrate is an important issue in the voltage readout pixel array. Wires and S/D diffusions are capacitively coupled to the substrate. When they change voltage, they induce currents in the substrate, which in turn cause voltage fluctuations in the substrate. The impact of substrate noise can be reduced by using guard rings around noise sources and receivers, multiple ground pins to reduce the effective bond wire inductance, and fully-differential analog circuit design. These techniques are difficult to apply within the pixel array itself, since substrate contacts are seldom present inside the pixel and the pixel output is inherently single-ended. The column lines are particularly troublesome because they are close

to adjacent pixel columns, undergo appreciable voltage swings, and extend for the entire length of the array. Crosstalk occurs when a voltage change on one column line induces a voltage change on another column line which does not settle before the signal is sampled.

Simplified plan and cross-sectional views of the pixel array are shown in Fig. 2-26. An accurate simulation would require a full three-dimensional simulation using the actual p-well and S/D diffusion doping profiles. Fortunately, a simpler simulation can indicate whether or not crosstalk is likely to be a problem. A column line and the underlying substrate act as an R-C circuit, where C is the capacitance between the column line and substrate, and R is the resistance between the silicon beneath the column line to the substrate contact. The exact field doping profile for the process is unknown, but is not expected to differ too much from that used in the MIT baseline twin-well CMOS process [39]. The MIT process uses epitaxial wafers; it is assumed that the process in which the imager chip is faabricated also uses epitaxial wafers. For the MIT CMOS process, the p-well doping is estimated to lie between about 10^{15} cm^{-3} and 3×10^{15} cm^{-3} between 0 μm and 6 μm into the silicon. For depths exceeding 6 μm, doping increases approximately exponentially with depth, up to the point where the silicon becomes degenerately doped. The doping exceeds 2 *times* 10^{16} cm^{-3} for depths greater than 6.5 μm; it is assumed that the substrate is effectively ground beyond this depth. The surface has a highly doped field implant, but it is irrelevant for this simulation.

The simulated structure is shown in Fig. 2-27. The gap between pixels in the same column is only 1.6 μm, which is small compared to the total pixel height of 22 μm. The structure is therefore approximately two-dimensional. The S/D diffusions are very shallow; they are only about 0.25 μm for the MIT 2 μm process, and are even shallower for a 0.8 μm process. Thus, their thickness is ignored. The p-well thickness is approximately 6.5 μm, based on the MIT process parameters. The substrate current flow is symmetric about the midpoint of the column line; therefore only the right half of the structure is simulated.

The substrate current was simulated with a 1 V bias applied between the top and bottom electrodes. Figure 2-28 is a plot of voltage along the silicon surface as a function of distance from the center of the column line. The net induced voltage on an n$^+$ photodiode would be found by averaging this voltage distribution over the length of the photodiode. Induced voltages clearly decay rapidly with distance from the disturbance: approximately one order of magnitude for every 10 μm. Given a pixel 22 μm wide, the voltage disturbance is significant only for the pixel columns closest to the column line. The conclusion is that electrical crosstalk due to the coupling between column lines is not a significant problem, given the assumed conditions.

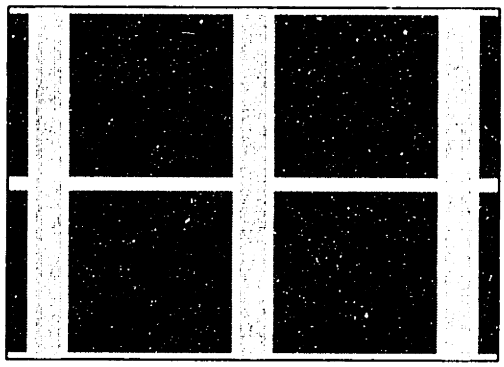


Figure 2-26: (a) Plan view of column line and pixels. (b) Cross-sectional view.

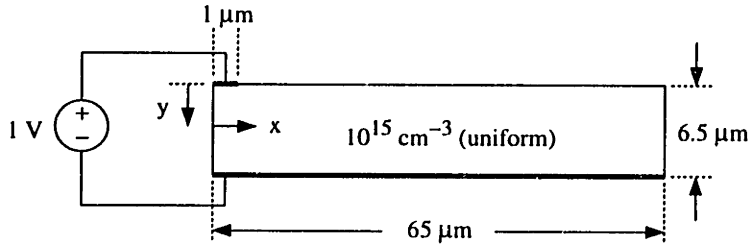


Figure 2-27: Simulation structure for crosstalk.

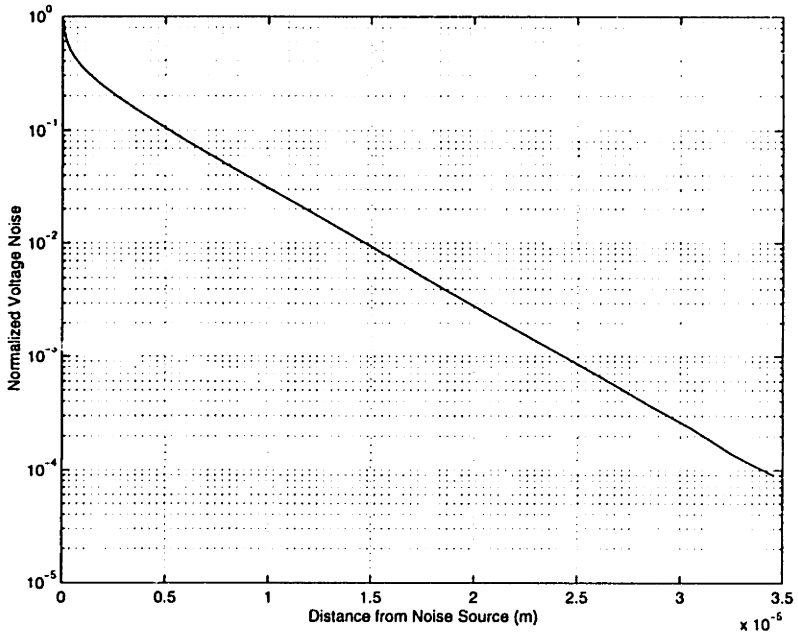


Figure 2-28: Attenuation of voltage disturbance along silicon surface.

Anti-blooming

The voltage readout pixel has the same anti-blooming options as the charge readout pixel. Automatic anti-blooming is provided by the reset device and the column line S/D diffusions.

2.4 Comparison of CCD and CMOS Imagers

Table 2-5 summarizes the performance specifications for an IL CCD, CMOS charge readout, and CMOS voltage readout pixel arrays. The CMOS current readout pixel is very similar to the voltage readout pixel, and is therefore not listed separately. Some entries depend on the particular process which is used.

Several processes can be used for imagers. The NMOS/CCD process is currently the most commonly used for this application. Depletion-mode NMOS devices are formed using the buried-channel implant; they are typically used in the output source follower chain. A CCD/CMOS process adds enhancement-mode NMOS and PMOS devices [40]. A digital CMOS process (D-CMOS) has enhancement-mode NMOS and PMOS devices, one poly layer, and no capacitors aside from those normally considered parasitic. A enhanced CMOS process (E-CMOS) has additional features to enhance its imaging capabilities, such as two poly layers

and tailored pixel doping profiles.

2.4.1 Quantum Efficiency

CCD and enhanced CMOS processes should have equivalent collection gain, since the pixel doping profile is similar. Digital CMOS processes have to use either the n-well or S/D diffusion. The n-well should have good collection gain for two reasons. The n-well is deep enough that most photocharge is generated inside the diffusion, where it cannot wander away. Also, the n-well doping is light enough so that the implant damage can be annealed out. The produces fewer traps and a lower likelihood that a generated electron-hole pair is recombined. The S/D diffusion is likely to have poor collection efficiency because the implant is shallow. Recombination is high for the shorter wavelengths, since they get absorbed in the silicon close to the silicon-oxide interface, where the trap density is highest.

Optical aperture is limited by the vertical shift register in the IL CCD, and by wires, transistors, and contacts in a CMOS array. The optical aperture for a photogate IL CCD array with a four-phase shift register cannot exceed 33% if the shift register has the same channel doping as the photogate. Otherwise, the shift register will overflow with a saturation-level charge packet. The optical aperture can be higher with a photodiode pixel, or when pixel capacity is limited by an overflow drain.

The charge readout pixel tends to have higher optical aperture than the voltage readout pixel since it has one fewer wires, two fewer transistors, and three fewer contacts.

2.4.2 Conversion Gain

Conversion gain is limited only by minimum design rules for the IL CCD and charge readout pixel arrays. In the simplest voltage readout pixel, the charge sense capacitance is roughly proportional to the size of the active area. Conversion gain can be greatly improved by having separate charge collection and charge sense nodes. Using separate regions allows a large, high capacitance diffusion to be used to collect a large amount of charge. The charge sense node is made very small so that it has low capacitance.

2.4.3 Random Noise

Photonic shot noise and dark current shot noise are common to all imagers. Mean square dark shot noise is proportional to dark current, so this component can be reduced by minimizing dark current. Digital CMOS processes are at a disadvantage because their depletion regions extend

to the silicon-oxide interface, a region with a high concentration of recombination-generation sites.

Digital CMOS processes also produce kTC noise during pixel reset. CCD's do not suffer from this because the charge is completely transfered out of the pixel at readout. With appropriate enhancements, a CMOS process can completely transfer charge out of the charge integration node, performing a noiseless reset. For example, an imaging array using this technique has achieved 42 e^- noise [41], which was dominated by the shot noise of the dark current. A digital CMOS process can approximate this using a small diffusion between the photogate and transfer gate.

The charge readout pixel also suffers from amplifier noise. This noise source is particularly important because of the large closed-loop gain of the output circuit.

2.4.4 Saturation Level

Saturation level is the product of pixel area and electron capacity per unit area, which in turn depends on optical aperture and doping profile. CCD pixels typically have saturation levels of 1000 - 3000 $e^-/\mu m^2$. Reported pixel saturation levels for voltage output CMOS pixels include 3750 $e^-/\mu m^2$ [41] and 7500 $e^-/\mu m^2$ [42], however, imager saturation occurs at lower illumination than pixel charge saturation. The charge saturation level is proportional to optical aperture, so the current output pixel should have a charge saturation level similar to that of the voltage output pixel, and the charge output pixel should have a charge saturation level approximately double that of the voltage readout pixel.

2.4.5 FPN

Mismatches in optical aperture and dark current are sources of FPN in all pixel arrays. Dark current magnitude, and therefore mismatch, can be reduced by appropriate pixel doping profiles. Dark current FPN is therefore likely to be higher in a digital CMOS process than either a CCD process or an enhanced CMOS process.

Readout capacitance mismatch is only present when multiple charge sense nodes are used. A CCD or charge readout array can use a single output circuit for the entire array, but the voltage readout array inherently uses a separate charge sense node for each pixel.

The voltage readout pixel also has significant FPN due to mismatch in source follower gain and mismatch in the reset transistor gate-source overlap capacitance. The charge readout pixel is insensitive to overlap capacitance mismatch because the same charge is injected into the pixel

each time it is reset.

The best FPN performance is produced by a CCD array and charge readout CMOS array with a single output circuit for the entire array. They have no sources of FPN beyond those common to all pixel arrays. FPN for the CCD array is lower than the FPN for a charge readout array built in a digital CMOS process since the digital CMOS process has higher dark current.

FPN is next best for the voltage readout pixel. Compared to the charge readout pixel, the voltage readout pixel has additional FPN due to the reset transistor and charge sense capacitance, which are different for each pixel. The conversion between voltage on the charge sense node and output voltage is given by the source follower gain, which is not highly sensitive to bias current or source follower device aspect ratio.

FPN is highest for the current output pixel. It has all the FPN sources of the voltage readout pixel. In addition, it has gain FPN from mismatches in the bias current device (M4 in Fig. 2-19), the aspect ratio of the in-pixel active device (M1 in Fig. 2-19) and the output device aspect ratio (M5 in Fig. 2-19).

2.4.6 Image Lag

The CCD pixel array does not suffer from lag due to the complete charge transfer property. A charge readout pixel leaves a small proportion of its charge in the pixel at readout, resulting in slight image lag. A voltage readout pixel has significant image lag only if the pixel uses a soft reset. The charge sense node is reset so quickly during a hard reset that significant image lag is unlikely.

2.4.7 Crosstalk

All pixel arrays have optical crosstalk between adjacent pixels, which is not usually a serious problem. CCD and charge readout pixel arrays suffer from smear, which is optical crosstalk from pixels into the adjacent shift register. Voltage readout pixel arrays use a low-impedance column line which is insensitive to leakage photocurrent, and therefore do not experience smear.

The CCD is immune to electrical crosstalk, up to the point where the charge is converted to voltage. The charge readout and voltage readout CMOS pixel arrays are potentially sensitive to crosstalk due to the lack of substrate contacts inside the pixel array. Electrical crosstalk caused by analog or digital circuits elsewhere on the die is a concern; return current paths through the substrate should be routed outside the pixel array. There are no inherent crosstalk problems in a CMOS pixel array; the substrate disturbance caused by a column line does not extend far

Table 2-5: Comparison of CCD and CMOS Imaging.

	Interline CCD	Charge Readout	Voltage Readout
coll. gain	high	fair (D-CMOS) high (E-CMOS)	fair (D-CMOS) high (E-CMOS)
optical aperture	< 33% (photogate)	2 W, 1 T, 2 C	3 W, 3 T, 5 C
conv. gain	$1/C_{sense}$	$1/C_{fb}$	$1/C_{sense}$
random noise	photonic noise dark current shot	photonic noise dark current shot kTC reset (D-CMOS) amplifier	photonic noise dark current shot kTC reset (D-CMOS)
FPN	optical aperture dark current C_{sense} (mult. output)	optical aperture dark current C_{fb} (mult. output)	optical aperture dark current C_{sense} SF gain reset device C_{ol}
saturation	1000 - 3000 $e^-/\mu m$		
anti-blooming	VOD	LOG (D-CMOS) VOD (E-CMOS)	LOG (D-CMOS) VOD (E-CMOS)
lag	none	incomplete PD reset	incomplete PD reset
optical crosstalk	adjacent pixels pixel to shift register	adjacent pixels pixel to column line	adjacent pixels
electrical crosstalk	none	minimal (column-to-column)	minimal (column-to-column)
system integration	CCD/NMOS: low CCD/CMOS: high	high	high

spatially or temporally.

2.4.8 Anti-Blooming

Lateral overflow gates are available in any of the architectures. An effective vertical overflow drain (VOD) can be made in a typical CCD process or an enhanced CMOS process.

2.4.9 System Integration

It is impractical to build sophisticated analog or digital circuits in an NMOS/CCD process. Usually, the only circuit built in such a process is the source follower chain that buffers the output diffusion voltage. A CMOS/CCD process can offer the same devices available in a digital CMOS process, allowing the integration of gain correction, color balance, gamma correction, analog-to-digital conversion, etc.

Both digital and enhanced CMOS processes allow high levels of system integration. Enhanced CMOS processes can provide good floating capacitors, which make switched-capacitor designs more practical.

2.5 Dynamic Range Expansion

One major limitation on imager performance is dynamic range. Imager dynamic range refers to the ratio of the highest detectable illumination to the lowest detectable illumination. If the dynamic range of the scene exceeds the dynamic range of the imager, portions of the scene will saturate the imager and appear either completely black or completely white. This is a common problem in outdoor scenes where the dynamic range may be as much as five orders of magnitude [16].

Pixel dynamic range is limited by noise on the low end and pixel saturation on the high end. Most CMOS pixels, like the CCD, are linear imaging devices. The stored charge is proportional to the irradiance, and the sensitivity of the imager is independent of illumination. In order to extend the upper end of useable irradiances, the sensitivity must be reduced. This does not increase the dynamic range since the lowest illumination which brings the signal above the noise floor is increased in the same proportion. What is desired is a pixel which has high sensitivity for low irradiance, and low sensitivity for high irradiance.

Two types of techniques have been developed to extend dynamic range. One technique uses a nonlinear imaging element, usually the gate-source voltage of a subthreshold MOSFET or the voltage across a forward-biased diode. The other technique combines information from several frames of a linear imager using multiple integration times. The specific techniques are discussed below.

2.5.1 Nonlinear Imaging Element

Antiblooming pixels have an inherent nonlinearity near the point at which the pixel saturates [43]. This allows for a simple type of compression using only a lateral overflow gate. Figure 2-29 shows one implementation of this idea. The blooming gate and imaging gate are held at fixed potentials. During the integration period, the transfer gate is held at a low potential to prevent electron flow into the output diffusion. At the end of the integration period, charge is read out in a manner similar to a CCD output. The output diffusion is reset by pulsing V_{RG} high and low, and then the transfer gate is pulsed high and low to allow the imaged charge to flow into the output diffusion. Reading out the charge also resets the imaging gate.

If the illumination is low, the charge level under the imaging gate is always substantially below the potential barrier under the blooming gate. All charge produced under the imaging gate remains there, and the integrated charge is proportional to the illumination. If the illumination is very high, the charge quickly reaches the level where all of the current flows across the

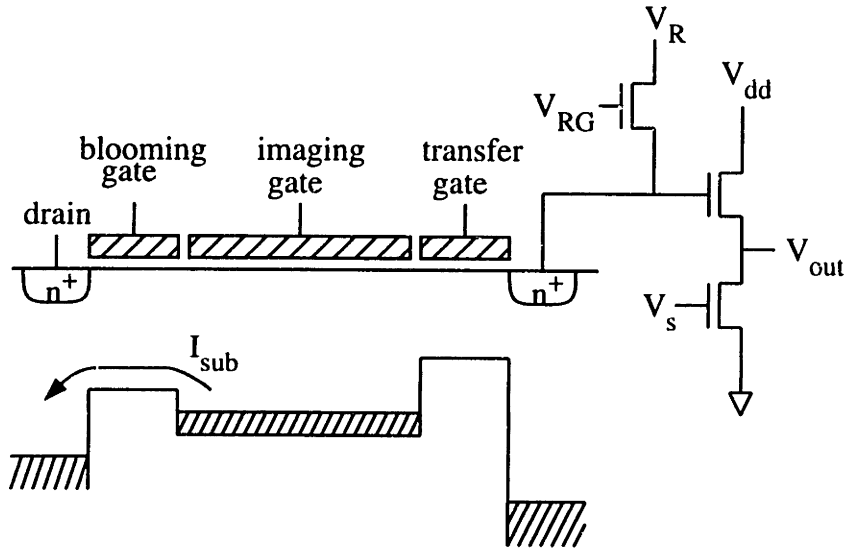


Figure 2-29: Compressing photodetector with fixed lateral overflow gate.

blooming drain. The integrated charge increases only slightly with illumination. In between these extremes, there is a range of illumination where the current divides between the channel capacitance and drain current. The total number of integrated electrons is derived in [43] as

$$N_{int} = -\frac{CV_{th}}{q} \ln\left[e^{-\Delta V_{sat}/V_{th}}(1 - e^{-I_P T_{int}/CV_{th}}) + e^{-I_P T_{int}/CV_{th}}\right]. \quad (2.58)$$

where C is the integration capacitance, V_{th} is the thermal voltage, ΔV_{sat} is the maximum voltage swing of the integration well channel potential, and I_P is the photocurrent.

This photodetector has low image lag, due to the complete reset operation. Compression is somewhat limited by the readout time; photocharge imaged during the time the transfer gate is high adds directly to the output charge. The longer the time required for transfer, the lower the achievable compression. Random noise will be limited to dark current and photonic shot noise, plus a contribution from the barrier limiting process, which will be discussed later. Resetting the output diffusion results in kT/C noise, but this is easily removed using CDS. The main drawback of this pixel is the inflexibility of the compression curve.

Another nonlinear imaging element uses an n^+ diffusion instead of an imaging gate, and ties the blooming gate to the diffusion [13]. The photodetector and associated potential diagram is shown in Fig. 2-30. The anti-blooming device M1 has a low threshold voltage to avoid cutoff; it always operates in subthreshold. As illumination increases, M1's source voltage decreases so that photocurrent and subthreshold current are in equilibrium. This is similar to the operation of the previous pixel, except that M1's gate is connected to its source rather than a fixed voltage.

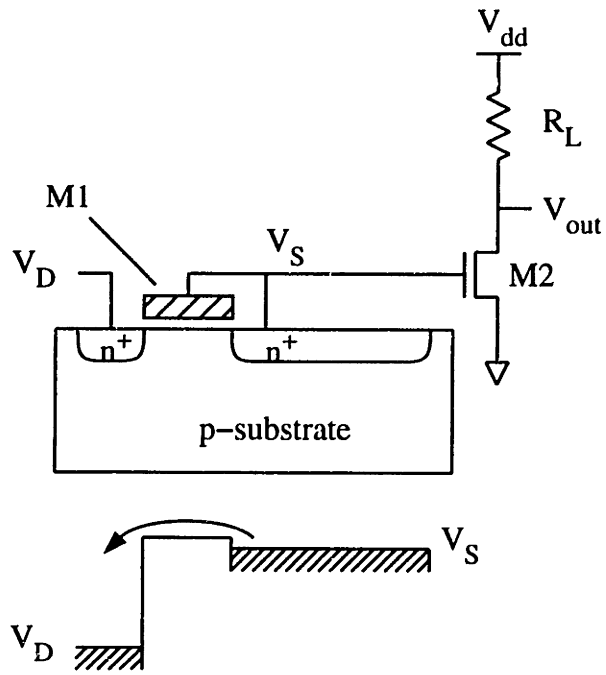


Figure 2-30: DALSA DYNASENSOR pixel. M1 is a depletion-mode device to allow current flow at zero gate-source voltage.

Optical aperture increases because there is no bias line to route, and sensitivity increases because a given change in illumination results in a larger source voltage change.

This pixel has a dynamic range of over seven orders of magnitude [13]. Random kTC noise in the pixel is estimated from different sources as either 46 e^- [44] or 142 e^- [45]. The noise equivalent power of a linear imager using these photodetectors is 10^{-10} W/cm^2 [44]. Image lag is likely to be a significant problem because the settling time constant is set by the large photodiode capacitance in parallel with a source impedance of a subthreshold MOSFET. Although similar to other CMOS pixel structures, this pixel cannot be implemented in a digital CMOS process since the anti-blooming device must be depletion-mode. There are no means for adjusting the compression characteristic.

An isolated photodiode exposed to light will become forward-biased so that the forward-bias current matches the photocurrent. The voltage is proportional to the logarithm of the irradiance, following a 60 mV/decade relationship. Mann uses a photodiode operating in this mode in the design of an adaptive photoreceptor [14]. Dark current is the sole source of fixed-pattern noise, since the relation between diode voltage and current is based on fundamental

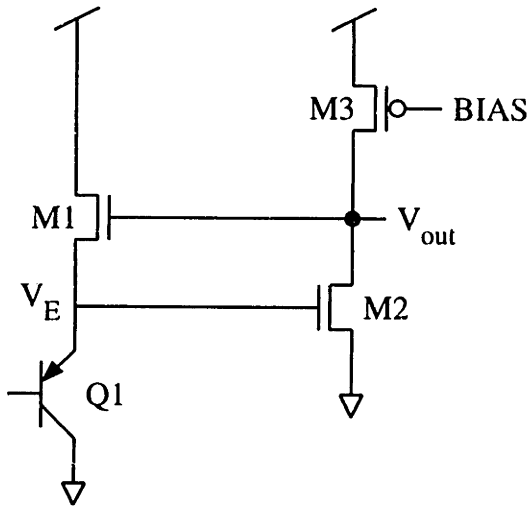


Figure 2-31: Silicon retina photoreceptor.

device physics. Random noise is similar to the reset noise for an integrating-mode photodiode.

$$v_{n,rms}^2 = \frac{i_{n,rms}^2}{g_d^2} = \frac{2q(I+I_d)}{g_d^2} \cdot \text{BW (Hz)} = \frac{1}{2}\frac{kT}{C_d}, \tag{2.59}$$

where C_d is the total photodiode capacitance, which comprises a depletion capacitance term and a diffusion capacitance term. The pixel output voltage has a large temperature coefficient.

The forward biased photodiode performs two functions, converting light to current and converting current to voltage. If these functions are separated, a linear element can be used to transduce the light, and a nonlinear element to convert from current to voltage. An adaptive photoreceptor using this approach is shown in Fig. 2-31 [15]. The open-base bipolar Q1 acts as a phototransistor. The base current is the photocurrent, which gets amplified by the β of the transistor. This amplified current passes through M1, setting $V_{gs,1}$ via the subthreshold current equation. M2 and M3 form an amplifier which keeps M1's emitter voltage approximately constant. The gate-source voltage $V_{gs,1}$ appears entirely at the gate of M1, which is also the output. The output voltage will be

$$V_{out} = nV_T \ln\left(\frac{I_p + I_d + I_S}{I_S}\right) + V_E, \tag{2.60}$$

where I_p is the photocurrent and I_d is the pixel dark current.

The large number of transistors will tend to cause large random noise and increase pixel size. The actual random noise will depend on the capacitances present in the pixel, and force a trade-off between noise and image lag. Higher internal capacitances produce lower noise but greater image lag.

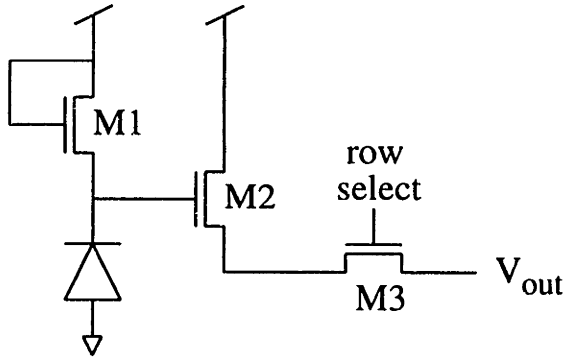


Figure 2-32: Subthreshold-mode MOS pixel.

Another variant using a reverse-biased photodiode is shown in Fig. 2-32 [17]. Here, the photocurrent passes directly through M1, which is in subthreshold. Fixed-pattern noise reported for a 256 × 256 pixel array was large; 60 mV for pixel FPN and 20 mV for column FPN. This could be improved by CDS, but CDS is difficult to implement due to the continuous-time nature of the pixel. Response is logarithmic over six orders of magnitude of illumination. Like the previously dicussed techniques, the compression is non-adjustable. Seger *et al.* have built and tested a similar prototype imager [16].

The CCD, which is normally a linear imaging device, can be turned into a nonlinear element by selectively transferring charge to the vertical shift register. Figure 2-33 [46] shows one scheme. At the start of the integration period, ϕ_{AB} is low, creating a potential barrier to electron flow into the sink. The photosite sampling gate clock ϕ_S is high. As photocharge accumulates in the pixel, it first flows under the storage gate since the potential is highest there. As additional charge accumulates, it spreads out under the storage gate, photosite sampling gate, and photogate. At the end of the integration period, the photosite sampling gate goes low, splitting the charge. The anti-blooming gate goes high, dumping the charge under the photogate to the sink. The charge under the storage gate is clocked through the transfer gate into the vertical shift register, where it is sent to an output circuit. Using one storage gate produces a two-slope compression curve, but the concept can be extended to multiple storage gates and compression curves with multiple sloped regions.

The CCD version of this compressive pixel should have insignificant image lag, due to the complete reset of the pixel. A properly-designed CMOS version will also have low image lag. The increase in dynamic range is the ratio of the total photogate area to the storage gate area, which is limited by design rules. Random noise will have the usual dark current and photonic shot noise components, plus an additional term from the partition noise of the photosite sampling

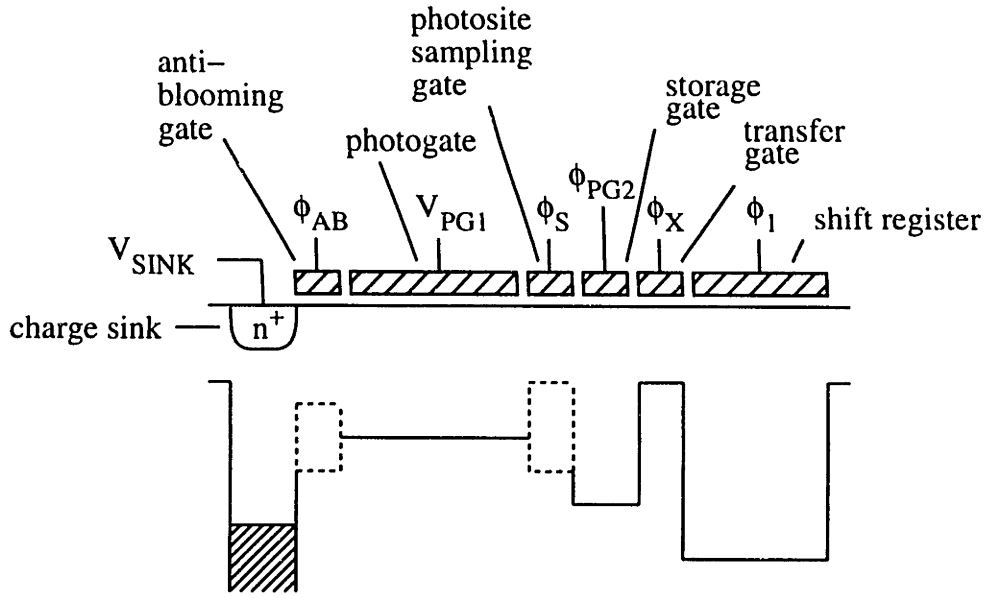


Figure 2-33: Area-based nonlinear CCD pixel.

gate. Fixed-pattern noise is likely to be large due to variation in the channel potentials under the photogate and storage gate(s), and asymmetry in the charge partition gate. This FPN is difficult to correct.

Another variant of a CCD compressive pixel is shown in Fig. 2-34 [18]. Photocharge is integrated under the photogates PCK and SCK. For low illumination, the channel potential under PCK and SCK is greater than the potential under the BIAS gate, and all the charge stays under the photogates. For high illumination, gates SC1, SC2, and BIAS implement a switched-capacitor resistor which drains off a fraction of the photocharge in excess of the knee charge Q_{knee}. The knee charge is given by

$$Q_{knee} = A_{pg}C_{chan}(\phi_{chan,pg}(0) - \phi_{chan,bias}(0)), \tag{2.61}$$

where A_{pg} is the combined photogate area (PCK and SCK), $\phi_{chan,pg}(0)$ is the zero-charge channel potential under the photogates, and $\phi_{chan,bias}(0)$ is the zero-bias charge under the BIAS gate. This produces a two-slope compression curve where the knee point is adjusted by the potential on the bias gate and the second slope is controlled by the clocking frequency.

2.5.2 Multiple Integration Periods

The other major approach to enhanced dynamic range is the combination of information from multiple integration periods (or multiple iris settings). Shorter integration times allow imaging

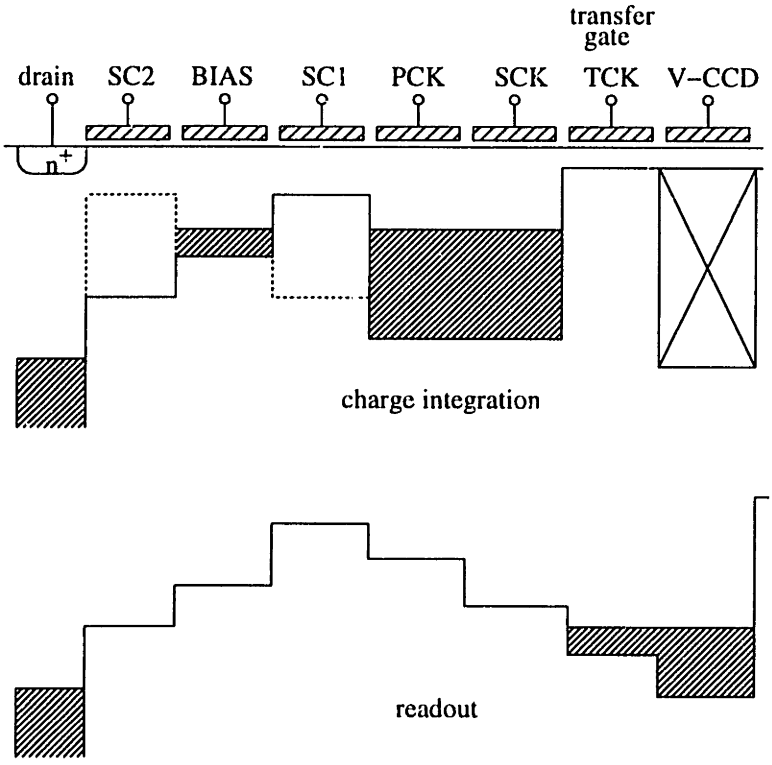


Figure 2-34: Nonlinear CCD pixel using switched-capacitor resistor.

of the brighter regions, and longer integration times allow imaging of the darker regions. Overall dynamic range is the product of the camera dynamic range and the ratio of longest to shortest integration times.

This operation can be performed by a system external to the imager. A camera control unit (CCU) [21] controls the integration period and downloads several frames at different integration times. The CCU combines this frame information, providing a transparent interface to the system designer. A CCU is a simple technique for upgrading an existing system, but limits the reduction of overall system size and power.

Another technique, the "Hyper-D range IL-CCD", uses the vertical blanking period to integrate a second frame [20]. The integration period of the second frame is user-adjustable and can be extremely short (10 μs). Charge packets from the two frames are transfered to separate gates in the vertical shift register and shifted out. At the output, an external circuit decides for each pixel whether or not that pixel saturated during the first frame (long integration time). If it did saturate, the effective pixel value is the pixel value from the second frame, multiplied by the ratio of the long integration time to the short integration time. Otherwise, the effective pixel value is simply the pixel value from the first frame.

For example, assume that the long integration period is 1/30 s, the short integration period is 1/300 s, and the pixel saturation level is 100,000 electrons. If the illumination produces a photocurrent of 3×10^5 e^-/s, the charge integrated over 1/30 s is 10,000 e^-. The pixel does not saturate, and the effective pixel value is 10,000 e^-. If the photocurrent is 6×10^6 e^-/s, the charge integrated over 1/30 s would be 200,000 e^-, except that the pixel saturates during the first frame, limiting the collected charge to 100,000 e^-. The charge integrated during the second frame is 20,000 e^-. This value is multiplied by the ratio of the integration times (10) to obtain an effective pixel value of 200,000 e^-.

Random noise, fixed-pattern noise, and image lag should all have values typical of an IL-CCD imager. The dynamic range is extended by the ratio of the integration times of the first and second frames, which is as much as 3300 given the specifications in Komobuchi's paper.

The drawbacks of the technique are the somewhat limited flexibility and the need for external signal processing. The technique is only practical for two-slope compression curves. The vertical shift register must simultaneously store n frames for an n-slope compression curve, and $n = 2$ is the practical limit.

It is possible to use a single integrating well with different effective integration times for light of different intensities, using a lateral overflow gate. One implementation is shown in Fig. 2-35 [19]. The pixel cross-section is shown at the top. The first potential diagram shows

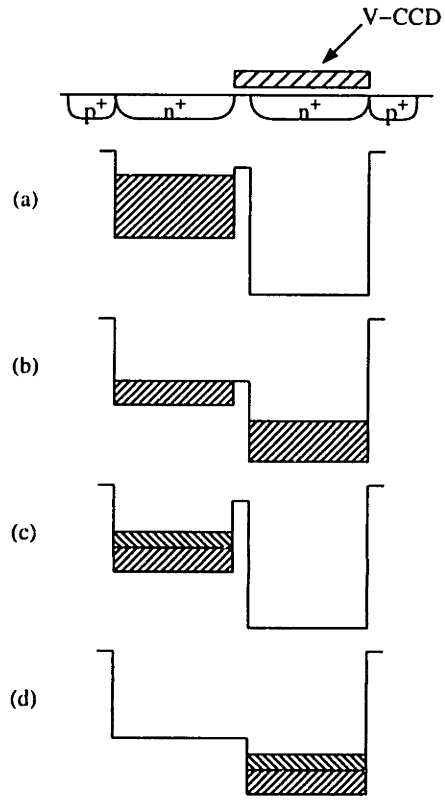


Figure 2-35: Timing-based nonlinear CCD pixel.

the situation near the end of the integration period. In Fig. 2-35(b), the transfer gate into the vertical shift register is partially on, clipping the charge level to a level determined by the V-CCD gate potential and the channel implant. If the irradiance is low, clipping does not remove any charge. Figure 2-35(c) shows the potential diagram at the end of the integration period. An additional amount of charge is integrated during the vertical blanking period, but the amount of extra charge is limited by the short integration time. In effect, the additional charge is integrated with a very short integration time. In Fig. 2-35(d), the charge is completely transferred into the vertical shift register. A two-slope compression curve is realized.

This scheme has many advantages over the other schemes mentioned in this chapter. Like the other CCD-based implementations, this pixel does not suffer from image lag. No additional area is required; the algorithm affects only the manner in which the transfer gate and vertical shift register operate. Threshold voltage mismatch in the channel leads to fixed-pattern noise, and random noise results mainly from the clipping operation. Only one packet is sent to the output circuit, requiring no change in the design of the vertical shift register. Although not

mentioned in the original paper, this technique can be extended to produce a multiple-slope transfer characteristic. This scheme is very similar to the scheme which was eventually chosen for the wide dynamic range imager described in this thesis.

2.6 Comparison of Compressive Imaging Techniques

The various techniques for compressive imaging are summarized in Table 2-6. Some general observations can be made. Most of these techniques are not programmable. The CCU is programmable, but does not allow a high level of system integration. The timing-based non-linear CCD imager is also programmable, but only to a limited extent.

The techniques which perform logarithmic compression tend to have very high dynamic range. The others apply two-slope compressive characteristics which extend dynamic range by only a modest factor. Dynamic range expansion in the area-based non-linear CCD is limited by photogate size, since the storage gate size is fixed by minimum design rules. High dynamic range is possible only with a very large pixel. The Hyper-D range IL-CCD and timing-based non-linear CCD have dynamic ranges fixed by the ratio of time intervals. The smaller time interval was the vertical blanking time, but a shorter interval could in principle be used. Ultimately, these CCD's are limited by smear.

The pixel chosen for this thesis is a variant of the timing-based non-linear CCD. The technique is extended to allow a programmable multiple-sloped compression curve, and implemented in a CMOS APS array to avoid the smear limitation.

2.7 Comparison of On-Chip ADC Architectures

Several converter architectures have been used for imaging arrays. The converter choice was based on the signal format, accuracy, sampling rate, and area. The signal format is a sequence of voltage steps as shown in Fig. 2-36, each step representing the output of one pixel. The required accuracy is 10 bits, although lower accuracy may be acceptable in some applications. The sampling rate depends on the number of channels. The product of the channel count and the converter sampling rate must equal the total number of samples per second, which is approximately 66 MS/s at a 1 KHz frame rate. This frame rate was one of the goals for the imager, and would be useful for certain machine vision tasks such as automated inspection. Area is the product of channel count and individual converter area. Area should be as small as possible.

Table 2-6: Comparison of compressive imaging techniques.

	Pixel Size	Programmable	Dynamic Range	Compression
pixel with anti-blooming	VOD or LOD required	no	10^8	linear (low light) log (high light)
DYNASENSOR	approx. same as std. active pixel	no	$> 10^7$	logarithmic
silicon retina photoreceptor	1 extra W + 1 extra T	no	$> 10^4$	logarithmic
forward-biased photodiode	no change	no	$> 10^6$	logarithmic
sub-threshold MOSFET	same as voltage readout	no	10^6	logarithmic
area-based non-linear CCD	3 extra gates + drain diff	no	$\times A_{pg}/A_{sg}$	2-slope
camera control unit (CCU)	off-chip memory and processor	yes	$> 10^4$	N.A.
Hyper-D range IL-CCD	no change	no	$\times 20$	2-slope
timing-based non-linear CCD	no change	knee point	$\times 6.8$	2-slope

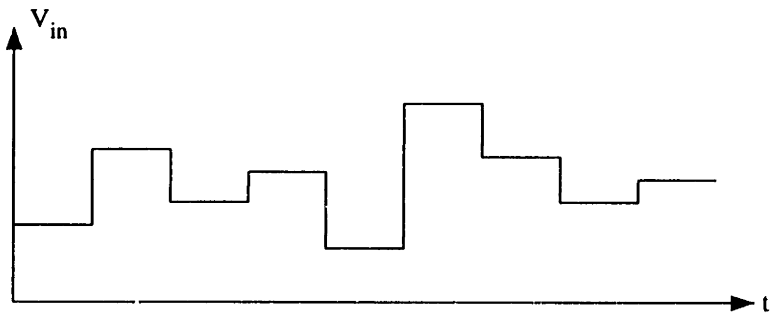


Figure 2-36: Typical signal waveform from pixel array.

2.7.1 Delta-Sigma Converter

The delta-sigma converter is used in high accuracy, low sampling rate applications. Accuracy is usually from 12 to 20 bits, and typical input signal bandwidth is in the several kilohertz region.

The delta-sigma converter is typically implemented as a switched-capacitor circuit. An L^{th}-order converter requires L integrators and one comparator. If C is the number of channels, the total area is

$$A_{\Delta-\Sigma} = L \cdot C \cdot A_{int}, \tag{2.62}$$

where A_{int} is the area of one integrator stage.

The total sampling rate is the product of the Nyquist sampling rate and the oversampling ratio. Assuming a one-pole anti-aliasing filter and 10-bit settling, the filter bandwidth must be at least

$$\text{BW} = \ln(2^{-10})\frac{66\,\text{MHz}}{C} \approx \frac{462}{C}\text{MHz} \tag{2.63}$$

in order for the filtered signal to settle adequately close to the true step value. The required Nyquist sampling rate is twice this bandwidth.

The oversampling rate (OSR) is determined by the allowable in-band quantization noise, which depends on the required conversion accuracy. The signal-to-noise ratio (SNR) at maximum input signal amplitude must be at least 2^{10} for 10-bit accuracy. The input signal cannot exceed the separation Δ between the two quantization levels for the converter to be stable. The in-band quantization noise is [47]

$$n_0 = e_{rms}\frac{\pi^L}{\sqrt{2L+1}}(\text{OSR})^{-(L+0.5)}, \tag{2.64}$$

where the total quantization noise is

$$e_{rms} = \frac{\Delta}{\sqrt{12}}. \tag{2.65}$$

Combining all of the preceeding equations, the required oversampling ratio is

$$\text{OSR} = \frac{\Delta}{n_0} = \left(\frac{2^{10}\pi^L}{\sqrt{12(2L+1)}}\right)^{1/(L+0.5)}. \tag{2.66}$$

The total sampling rate is

$$f_s = f_{Nyquist} \cdot \text{OSR} = 2\frac{462}{C}\left(\frac{2^{10}\pi^L}{\sqrt{12(2L+1)}}\right)^{1/(L+0.5)}\text{MHz} \tag{2.67}$$

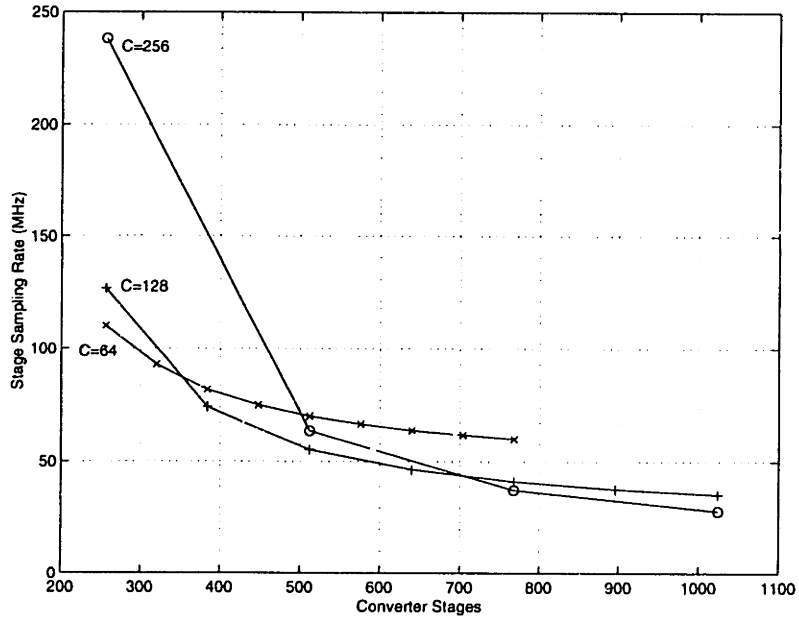


Figure 2-37: Sampling rate in oversampled $\Delta - \Sigma$ converter.

Figure 2-37 plots sampling rate as a function of area, expressed in terms of SC integrator stages. For any reasonable area, the SC integrator must operate above 25 MHz. As shown below, this sampling rate is much higher than that required for an algorithmic converter using a similar number of stages.

2.7.2 Flash Converter

Flash converters are used in low accuracy, very high speed applications. Typical accuracy is six bits, and typical conversion rates are in the tens to hundreds of megasamples per second range. Required area and power increase exponentially with accuracy; a 10-bit flash converter would require 1024 comparators, each offset cancelled to within about 1 mV. The large area and power required for these comparators is likely to make the flash architecture impractical.

2.7.3 Dual-Slope Converter

Dual-slope converters, like delta-sigma converters, are used in high accuracy, low sampling rate applications. It is an attractive option, since the digital counter and reference voltage ramp can be shared by all converters on a chip. Each converter requires only an RC integrator, comparator, and digital latch. Given the small area and low conversion rate, it is most practical

to use a separate channel for each column. The drawback is the very high speed operation of the global counter. To perform a 10-bit conversion in 4 μs, the count rate must exceed

$$f_{count,min} = \frac{2^{10}}{4\mu s}. \tag{2.68}$$

In other words, all converters must recieve the digital word in under 4 ns. Controlling skew is likely to be extremely difficult.

2.7.4 Algorithmic Converters

An algorithmic converter is a series of stages, each of which performs a mathematical operation on the input voltage and produces one bit and an output voltage. The output voltage, or residue, is the input for the following stage. A typical operation is to double the input and either add or subtract a reference voltage, depending on whether the input is positive or negative. The generated bit is 1 if positive, 0 if negative.

Algorithmic converters can be subdivided into pipeline and cyclic architectures. In the pipeline converter, separate SC integrators are used to implement each stage. In a cyclic converter, the same SC integrators are reused to generate subsequent bits. For 10-bit output, 10 stages are used in the pipeline converter. The cyclic converter can use any number of stages, but two stages were found to be most convenient in the actual design, and will be assumed for comparison. A typical pipeline converter operates at a 1-20 MS/s sampling rate, with an accuracy of about 10-12 bits uncalibrated, or 15-16 bits with digital calibration. A typical cyclic converter operates at sampling rates on the order of 10-100 KS/s, with an accuracy of 8-14 bits.

Sampling rate and area are inversely related for an algorithmic converter. Figure 2-38 shows this relation assuming that the SC integrator size is constant. Because of the similarity of these designs, their sampling rate – area curves overlap except at very small and very large area. However, neither of those regions are of practical interest.

In the small area limit, area is constrained to be less than the area of 10 comparators and 10 SC integrators. Only the cyclic architecture can be used. There is no realistic reason for the area to be so constrained, since 10 comparators and 10 SC integrators take up only 0.42 mm^2 in a 0.8 μm process, using parameters from the actual design. This area is insignificant compared to the pixel array size.

In the large area limit, there is more than enough area for every column to have its own cyclic converter. This area is greater than the area of 512 comparators and SC integrators, which is 21.4 mm^2, again using parameters from the actual prototype imager. This is over one-fourth of the total die area, which is undesirably high.

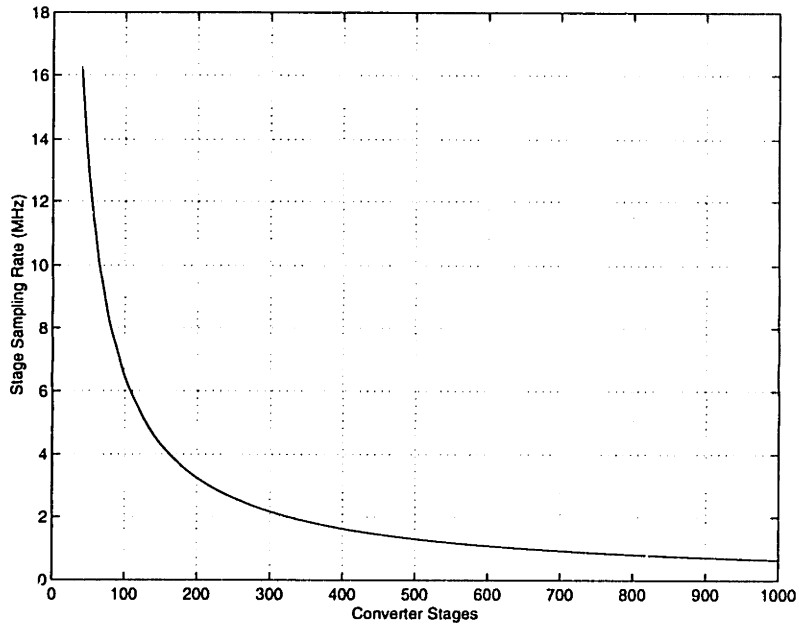


Figure 2-38: Stage sampling rate for algorithmic converter. Graph assumes 1000 frames/s conversion rate on a 256 × 256 pixel array.

The conclusion is that the cyclic and pipeline converters trade off roughly equally over the range of usable converter area. Since the cyclic converter has no major drawbacks, and consumes far less power than an oversampled delta-sigma converter, it was chosen for the prototype chip.

Chapter 3

The Wide Dynamic Range Pixel

3.1 Introduction

The wide dynamic range pixel used in this thesis uses a lateral overflow gate whose potential changes over the integration period to produce a non-linear compression curve [48, 49]. The implementation in this thesis is a CMOS voltage readout pixel, with CDS applied to reduce offsets. A plan view of the pixel is shown in Fig. 3-2, and a cross-section is shown in Fig. 3-1. An equivalent circuit is shown in Fig. 3-3.

The lateral overflow gate M4 is used to increase the dynamic range of the pixel. Over the integration period, the voltage on this gate decreases in steps; depending on the size, number, and timing of these steps, any compressive characteristic can be approximated. The charge spill gate M3 increases the sensitivity of the pixel. It acts as a common gate amplifier; photocurrent flows into the low impedance source node and is discharged into the high impedance drain. This allows charge collected by the large photodiode to be sensed by the small charge collection diffusion. The source follower M1 buffers the pixel from the large column line capacitance. The row select device M2 connects the source follower output to the column line when the row is read out.

The integration period starts with the lateral overflow gate and sense diffusion at high potential. M3 acts as a common-gate amplifier. Charge generated in the large photodiode diffuses across M3 into the sense diffusion, lowering its potential. M4 applies a time-varying potential barrier to electron flow. At the end of the integration period, M2 turns on, allowing a bias current to flow through the source follower device. After a suitable period of time to allow the voltage on the column line to settle, the pixel output voltage is sampled by the CDS

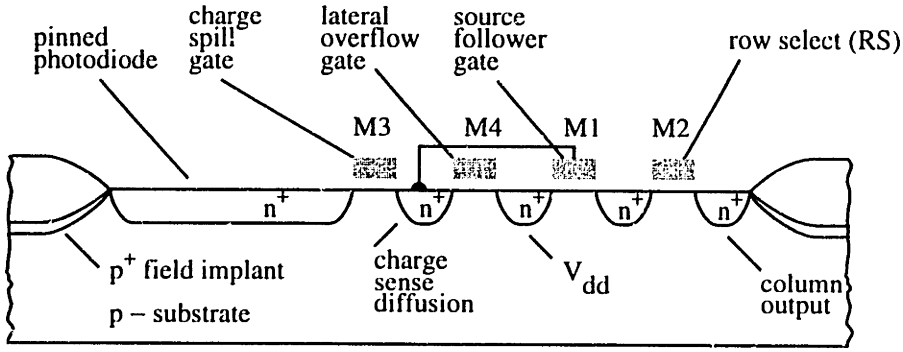


Figure 3-1: Pixel cross-section.

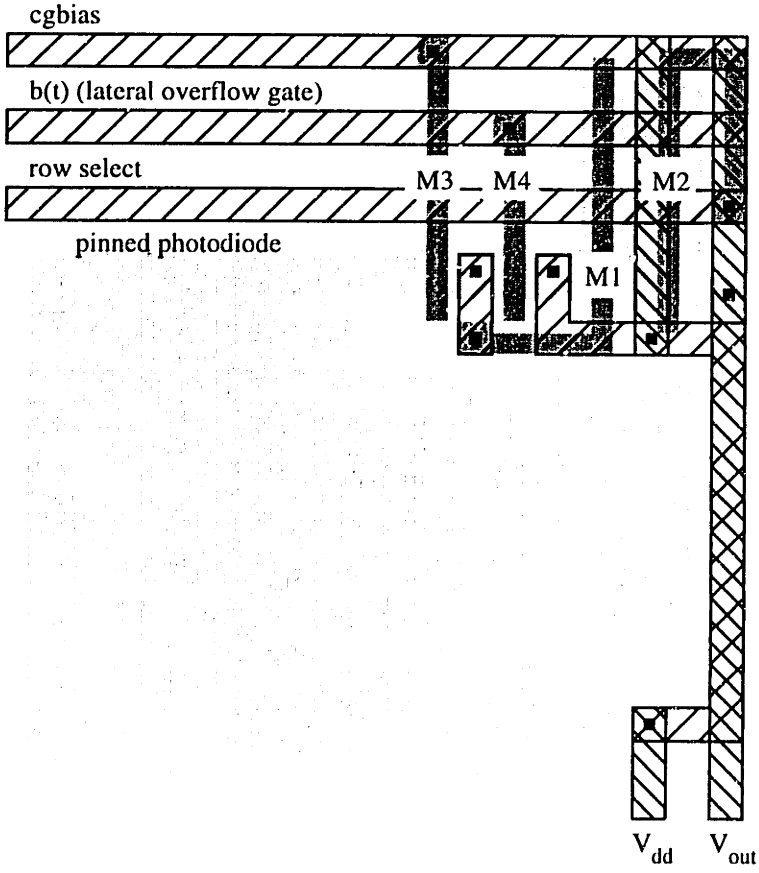


Figure 3-2: Plan view of wide dynamic range pixel.

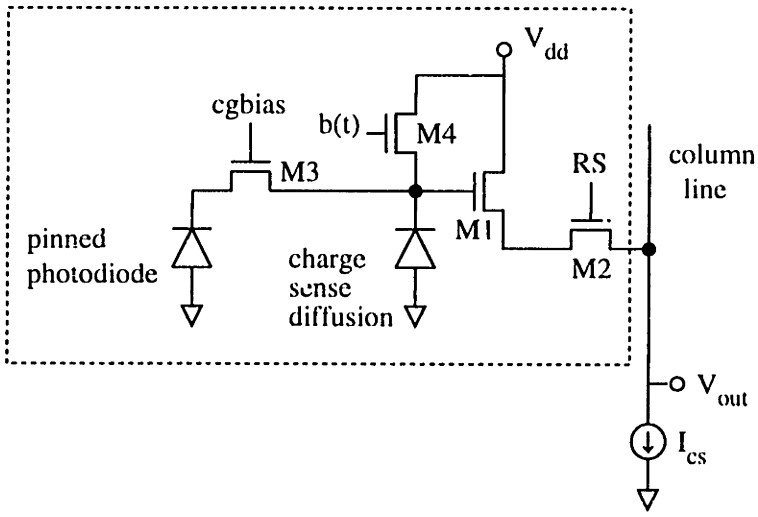


Figure 3-3: Equivalent circuit for the wide dynamic range pixel. M1 is the source follower device, M2 is the row select device, M3 is the charge spill device, and M4 is the lateral overflow gate. The dashed box shows the portion inside the pixel.

circuit. Next, M4's gate is pulled to its maximum voltage, resetting the pixel. After allowing the column line to settle, the pixel output voltage voltage is sampled again. M4's gate drops to the first step potential, and the integration period begins again.

Although it is not immediately obvious, it is possible to obtain any desired compressive characteristic by appropriately choosing the waveform applied to the lateral overflow gate. This chapter proceeds by analyzing the relationship between the barrier waveform $b(t)$ and the compression curve $Q(I)$ for the continuous-time case. Since it is difficult to generate and apply continuous-time waveforms to the imaging array, the continuous-time barrier waveform is approximated with a stepped waveform. The relationships between $b(t)$ and $Q(I)$ are analyzed for that case. The effects of nonidealities are taken into account. Finally, the pixel random noise and fixed-pattern noise are estimated for the stepped-barrier case.

3.2 Barrier Waveform and Compression Curve: Continuous-Time

3.2.1 Physical Modeling

The barrier is modeled with an ideal diode characteristic. Barrier current is zero when the charge integration curve is below the barrier, and infinite when the charge integration curve is

above the barrier. The barrier waveform $b(t)$ and the compression curve $Q(I)$ are plotted as increasing functions, although the actual photodiode voltage drops over time.

Figure 3-4 shows a typical barrier waveform and associated compression curve. The compression curve has three regions, defined by the photocurrents I_{min} and I_{max}.

For $I < I_{min}$, the charge integration curve is always below the barrier, and all photocharge stays in the pixel. The barrier is constrained to have a nonzero minimum value fQ_{max}, where Q_{max} is the well capacity and f is positive but otherwise unconstrained, so there will always some points on the compression curve corresponding to $I < I_{min}$. The final integrated charge is directly proportional to the photocurrent.

For $I_{min} < I < I_{max}$, the charge integration curve intersects the barrier curve at least once. During this time, the charge integration curve is limited to the barrier height. The barrier waveform slope is increasing, and eventually its slope exceeds the slope of the charge integration curve. At this point, the curves diverge and the charge is free to integrate for the remainder of the integration period.

For $I > I_{max}$, the photocurrent is large enough to saturate the charge sense diffusion before the end of the integration period. The maximum available integration time is T; its value is limited by the frame rate requirement. For example, for a 30 Hz frame rate, T cannot exceed 33 ms.

For analysis, it is sometimes useful to separate the total integrated charge into a barrier-limited portion Q_{bl} and a freely accumulated portion Q_{free}. The barrier-limited portion refers to the integrated charge at the time when the charge integration curve and barrier curve separate. The freely accumulated portion is the charge which is acquired after that time. If the charge integration curve and barrier curve never meet, all charge is freely accumulated charge.

The barrier waveform shown in Fig. 3-4 may be slightly altered to reset the pixel after readout. A CCD implementation requires no alteration of the barrier waveform; reset is performed automatically during readout. A voltage or current output pixel usually requires a reset transistor in the pixel. This could be a separate device, but it is more efficient to use the barrier transistor for reset. This requires that the barrier momentarily drop to the zero charge level between frames. Figure 3-5 compares altered and unaltered barrier waveforms. In the remainder of this chapter, the waveform will be drawn as in Fig. 3-5(a), but the actual waveform may be slightly different.

The relationship between the barrier waveform and compression curve has a "forward" problem and a "reverse" problem. In the forward problem, the goal is to find the compression curve for a given barrier waveform. In the reverse problem, the goal is to derive a barrier

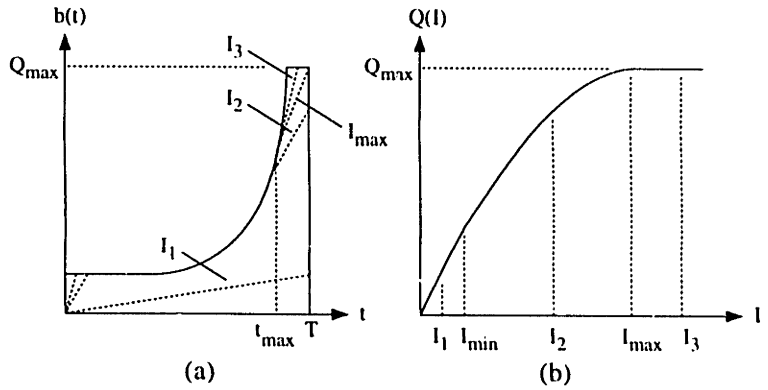


Figure 3-4: (a) Typical barrier waveform. I_1, I_2, and I_3 correspond to illumination levels falling in each of the three regions. (b) Compression curve for barrier waveform in (a).

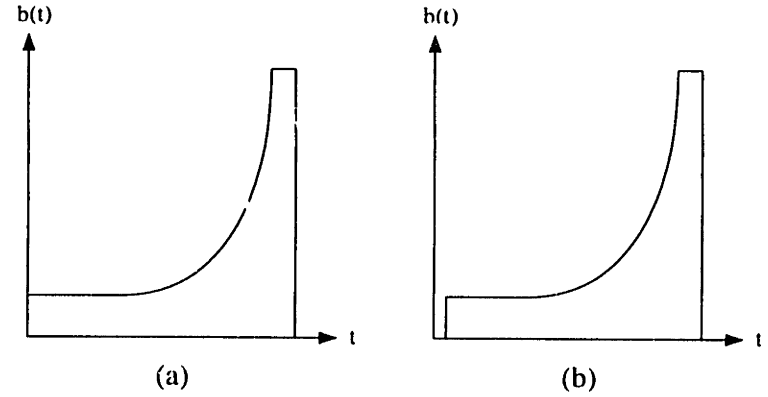


Figure 3-5: (a) Original barrier waveform. (b) Barrier waveform modified to perform pixel reset.

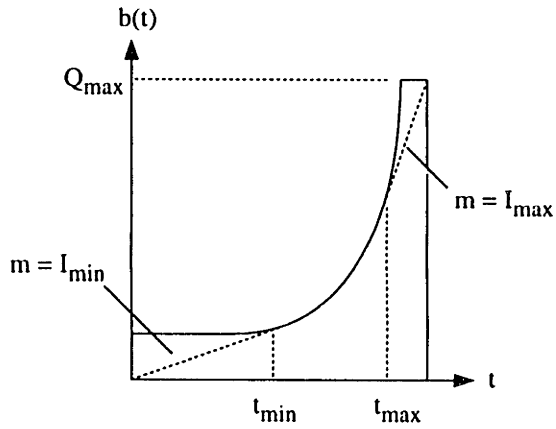


Figure 3-6: Determination of breakpoints I_{min} and I_{max}.

waveform that produces the desired compression curve.

3.2.2 Forward Problem

To solve the forward problem, the breakpoints I_{min} and I_{max} must be determined. The breakpoint I_{min} can be found conceptually by taking a straight line passing through the origin and rotating it counter-clockwise until it just touches the barrier waveform. The slope of this line defines I_{min}. This is shown in Fig. 3-6. If an equation for $b(t)$ is known, then I_{min} can be determined analytically.

$$I_{min} = b'(t_{min}) = \frac{b(t_{min})}{t_{min}}. \tag{3.1}$$

If this equation has no solution, then there is also no effective compression; the pixel behaves as a conventional linear imager.

Likewise, I_{max} can be found conceptually by taking a straight line passing through (T, Q_{max}) and rotating it clockwise until it just touches the barrier curve.

$$I_{max} = b'(t_{max}) = \frac{Q_{max} - b(t_{max})}{T - t_{max}}, \tag{3.2}$$

where t_{max} is the departure time corresponding to I_{max}. Aside from the degenerate case where $b(t) \geq Q_{max}$, this equation always has a solution.

For $I_{min} < I < I_{max}$, there is a departure time $t_d(I)$ at which the barrier and charge curves separate from each other. The total integrated charge is the charge stored in the well at that time, plus the charge added in the remainder of the integration period.

$$Q(I) = b(t_d) + I(T - t_d). \tag{3.3}$$

Since the slopes of the charge curve and barrier curve match at t_d.

$$I = b'(t)|_{t=t_d(I)}. \tag{3.4}$$

which can be inverted to give

$$t_d = b'^{-1}(I). \tag{3.5}$$

Combining these equations gives

$$Q(I) = b(t_d) + I(T - t_d)|_{t_d=b'^{-1}(I)}. \tag{3.6}$$

For example, assume a barrier waveform

$$b(t) = Q_{max}(ae^{t/\tau} + c). \tag{3.7}$$

Choose a and c so that

$$b(0) = fQ_{max} \tag{3.8}$$

$$b(t_{max}) = b(T) = Q_{max}. \tag{3.9}$$

Solving these gives

$$a = \frac{1-f}{e^{T/\tau} - 1} \tag{3.10}$$

$$c = \frac{fe^{T/\tau} - 1}{e^{T/\tau} - 1}. \tag{3.11}$$

The maximum detectable illumination is the given by the maximum slope of $b(t)$ which occurs at the end of the integration period.

$$I_{max} = b'(t)_{max} = Q_{max}(a/\tau)e^{T/\tau} \tag{3.12}$$

$$= \frac{Q_{max}}{\tau}\left[\frac{1-f}{1-e^{-T/\tau}}\right] \tag{3.13}$$

$$\approx \frac{Q_{max}}{\tau} \tag{3.14}$$

Dynamic range is increased by a factor of

$$\frac{I_{max}}{I_{max,nc}} = \frac{I_{max}}{Q_{max}/T} = \frac{T}{\tau}. \tag{3.15}$$

The departure times are

$$t_d(I) = \tau \ln(I\tau/aQ_{max}), \tag{3.16}$$

and the compression curve is

$$Q(I) = \begin{cases} TI & 0 \leq I \leq I_{min} \\ (Q_{max} - I_{max}\tau) + I\tau[1 - \ln(I/I_{max})] & I_{min} \leq I \leq I_{max} \\ Q_{max} & I \geq I_{max} \end{cases} \quad (3.17)$$

The dynamic range increase depends on the maximum allowable slope of $b(t)$. The integration time T is the reciprocal of the frame rate, typically 30 frames/sec. The maximum value of $b'(t)$ is inversely proportional to τ, so the dynamic range is proportional to the maximum slope of $b(t)$. If our system allows a rising edge slope of 1 V/100 ns to be accurately distributed over the array, the dynamic range expansion will be a factor of 400,000.

3.2.3 Reverse Problem

In this section, the barrier waveform $b(t)$ is derived given a desired compression curve $Q(I)$. The starting point is (3.3). Taking the derivative with respect to I gives

$$\begin{aligned} Q'(I) &= b'(t_d)t_d'(I) + (T - t_d) - It_d'(I) \\ &= (b'(t_d) - I)t_d'(I) + (T - t_d) \\ &= T - t_d, \end{aligned}$$

where the last equation results from 3.4. The departure time can be written as a function of I

$$t_d(I) = T - Q'(I), \quad (3.18)$$

which can be inverted to obtain I in terms of t_d

$$I(t_d) = (T - Q')^{-1}(t_d). \quad (3.19)$$

The barrier curve can be found by substituting (3.19) into (3.3) and substituting t for t_d.

$$b(t) = Q(I(t)) - I(t)(T - t), t_{d,min} \leq t \leq t_{d,max}. \quad (3.20)$$

It was assumed that the barrier curve never goes to zero. One consequence is that the compression curve must have a discontinuity in first derivative at I_{min}. The magnitude of this discontinuity determines $t_{d,min}$.

$$Q'(I_{min} - \epsilon) = T \quad (3.21)$$

$$Q'(I_{min} + \epsilon) = T - t_{d,min} \quad (3.22)$$

$$t_{d,min} = Q'(I_{min} - \epsilon) - Q'(I_{min} + \epsilon). \quad (3.23)$$

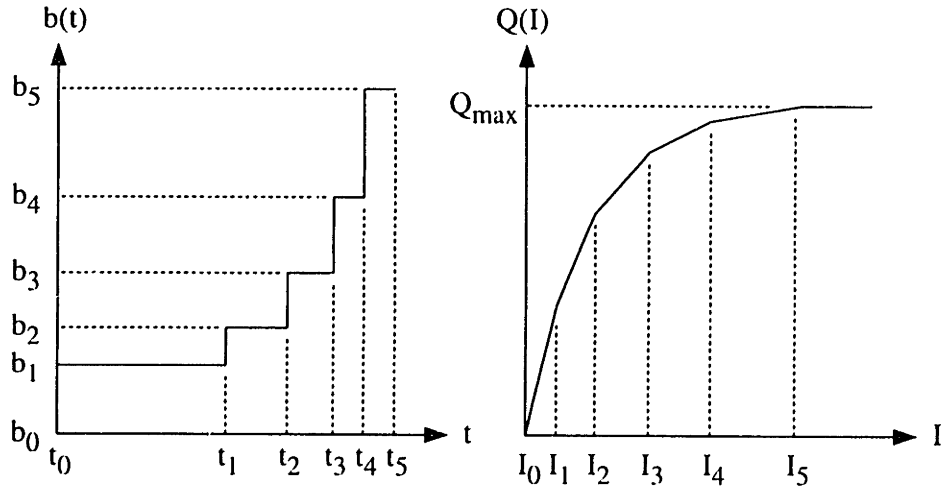


Figure 3-7: Stepped barrier curve and compression curve.

It is not possible, in general, to obtain a closed-form expression for $t_{d,max}$. An implicit relation between $t_{d,max}$ and $b(t_{d,max})$ can be found once $b(t)$ is calculated for $t_{d,min} \leq t \leq t_{d,max}$. Evaluating (3.20) at $t = t_{d,max}$ gives

$$b(t_{d,max}) = Q(I_{max}) - I_{max}(T - t_{d,max}) \quad (3.24)$$

$$= Q_{max} - I_{max}(T - t_{d,max}) \quad (3.25)$$

3.3 Barrier Waveform and Compression Curve: Stepped Barrier

It will be explained later how high dynamic range is incompatible with charge storage in the pixel. Each pixel must be read out immediately at the end of its integration period. Since it is impossible to read out all pixels at once, a staggered readout is required in which the integration period ends at a slightly different time for each row. Each row therefore requires a barrier waveform which is delayed from the barrier waveform of the preceeding row. As a practical matter, it is difficult to bring 256 continuous-time barrier waveforms onto the imager chip, or to generate multiple delayed versions of a continuous-time waveform.

The continuous-time barrier can be approximated with a stepped barrier. The stepped barrier is much easier to generate and stagger, as will be shown later. The compression curve obtained from a stepped barrier is an excellent approximation to the continuous-time compression curve, as shown in Fig. 3-7.

As before, the solutions to the forward and reverse problems can be obtained using the usual simplifying assumptions. The solutions to the forward and reverse problems are much simpler in the stepped-barrier case since the barrier and compression curves are described by only a limited set of numbers.

The barrier curve is specified by the N corners (t_i, b_i). As before, the barrier curve is constrained to lie above a certain minimum level. The highest barrier level corresponds to the saturation charge of the pixel in order to take advantage of the full pixel capacity.

$$b(t) = b_i > 0, \quad t_{i-1} \leq t < t_i, \quad 1 \leq i \leq N. \tag{3.26}$$

The reset level b_0 defines the zero charge level.

In order to use every barrier level, levels must be chosen so that the piecewise linear curve connecting the points $(t_i, b(t_i))$ is convex upward.

$$\frac{b_1}{t_1} < \frac{b_i - b_{i-1}}{t_i - t_{i-1}} < \frac{b_{i+1} - b_i}{t_{i+1} - t_i}, \quad 2 \leq i \leq N-1. \tag{3.27}$$

The compression curve is piecewise linear, with positive but decreasing slopes

$$s_i = Q'(I), \quad I_i \leq I \leq I_{i+1}, \quad 0 \leq i \leq N-1. \tag{3.28}$$

The compression curve always starts at zero, and limits at the saturation charge Q_{max} for $I \geq I_N$.

3.3.1 Forward Problem

In the forward problem, the barrier levels b_k and departure times t_k are known, and the compression curve breakpoints $(I_i, Q(I_i))$ must be calculated. Breakpoints occur when the charge integation curve departure point switches from $t_d = t_k$ to $t_d = t_{k+1}$. At this point, both t_k and t_{k+1} are departure points. The charge integration curve passes through (t_k, b_k) and (t_{k+1}, b_{k+1}).

$$I_i = \frac{b_i - b_{i-1}}{t_i - t_{i-1}}, \quad 1 \leq i \leq N \tag{3.29}$$

$$Q(I_i) = b_{i-1} + I_i(T - t_{i-1}), \quad 1 \leq i \leq N. \tag{3.30}$$

3.3.2 Reverse Problem

In the reverse problem, the breakpoints $(I_i, Q(I_i))$ in the compression curve are known, and the barrier levels b_k and departure times t_k are desired. The departure times are uniquely

determined by the slopes of the segments of the compression curve; once these are known, the barrier levels can be found by rearranging (3.30).

$$t_k = T - \frac{Q(I_{k+1}) - Q(I_k)}{I_{k+1} - I_k}, \quad 0 \leq k \leq N-1 \tag{3.31}$$

$$b_k = Q(I_{k+1}) - I_{k+1}(T - t_k), \quad 0 \leq k \leq N-1. \tag{3.32}$$

3.3.3 Comparison Between Stepped-Time Barrier and Continuous-Time Barrier

The number of steps required in the stepped-barrier approximation depends on the desired quality of fit to the continuous-time barrier. More steps produce a better fit, but additional steps require additional on-chip circuitry. This section will show a connection between "goodness of fit" and the number of steps.

The first issue is to figure out how the image degrades when only a few steps are used. Dynamic range *does not* degrade as long as there are at least two steps (aside from reset level). With two steps, one can use the entire integration period T_{int} for integrating low-level illumination and a very short integration time for very high-level illumination. The compression curves for the stepped-time (ST) and continuous-time (CT) barriers thus match at the low and high illumination limits. This situation is illustrated in Fig. 3-8(a).

The problem occurs in the region between the low-illumination and high-illumination limits. The sensitivity S, defined as the slope of the compression curve, can be very different for the ST and CT cases. There must be some range of illumination for which the ST compression curve is less sensitive than the desired CT compression curve. In Figs. 3-8(a) and (b), this range is $I_1 \leq I \leq I_{max}$. This appears in the image as reduced contrast for pixels with illumination between I_1 and I_{max}.

A reasonable standard for goodness of fit is the difference between the sensitivities for the ST and CT compression curves. Evenly spacing the steps along the sensitivity axis minimizes the maximum deviation between the ST and CT sensitivity curves. Two-step and three-step examples are shown in Figs. 3-8(a) and (b). For an n-step ST barrier curve, the maximum sensitivity difference is

$$\Delta S = \frac{S_{max} - S_{min}}{2n}, \tag{3.33}$$

where S_{max} is the maximum sensitivity of the CT compression curve and S_{min} is the minimum sensitivity. The maximum sensitivity deviation is inversely proportional to the number of steps.

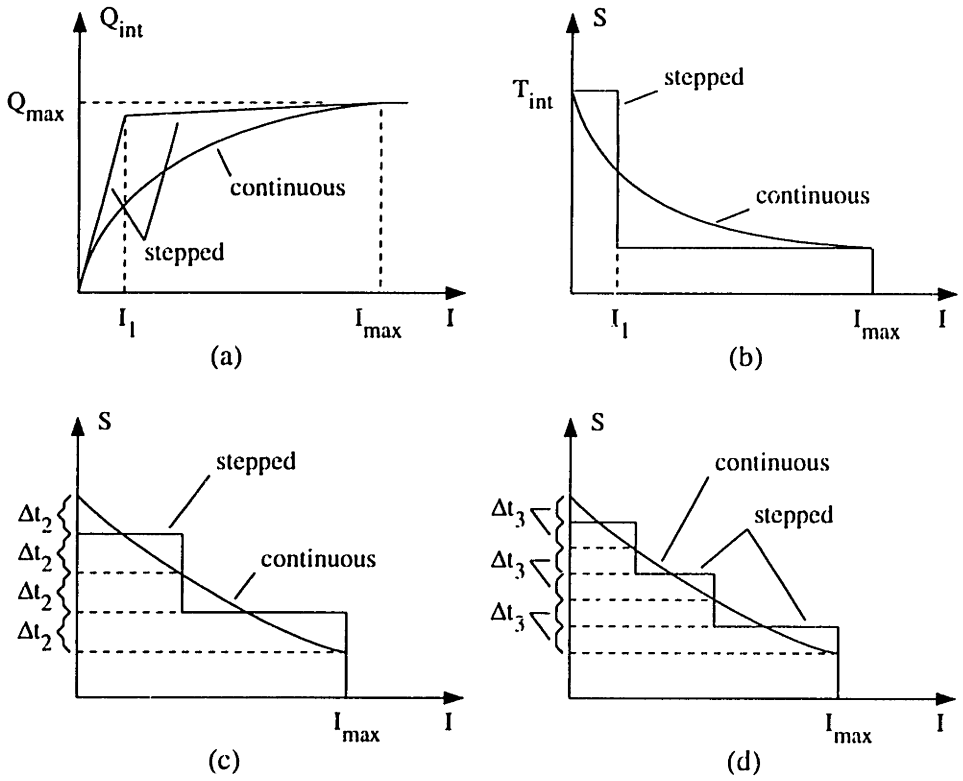


Figure 3-8: Comparison of stepped-time and continuous-time compression curves. (a) Two-step approximation to a continuous-time compression curve. The slopes are matched at $I = 0$ and $I = I_{max}$. (b) Sensitivity curves (first derivative) for compressioncurves in (a). (c) Two-step ST compression curve which minimizes maximum difference between CT and ST compression curves. (d) Three-step ST compression curve which minimizes maximum difference between CT and ST compression curves.

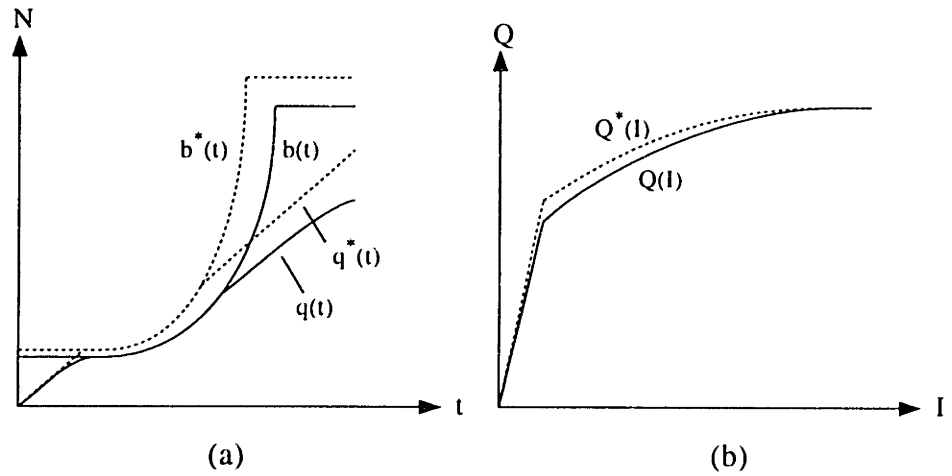


Figure 3-9: (a) Barrier and charge integration curves before and after nonlinearity compensation. (b) Charge compression curve before and after nonlinearity compensation.

3.4 Nonidealities

The physical situation is clearly more complicated than has been assumed in these derivations. It is necessary to understand how the simple physical model differs from reality, and how this affects the operation and performance of the algorithm.

3.4.1 Charge Integration Nonlinearity

It was assumed that the charge integration curve was a straight line for fixed illumination when it was not barrier limited. This is not precisely true since the charge sense capacitance increases as the photodiode voltage decreases. The charge integration curve $q(t)$ is warped in the vertical direction during linear integration periods, as shown by Fig. 3-9(a).

If the barrier and compression curves are compensated by a multiplying by $QE(0)/QE(Q)$, the linear integration regions are restored to straight lines. The compensated barrier curve is $b^*(t)$, and it produces a compression curve $Q^*(I)$, where $b^*(t)$ and $Q^*(I)$ are related to $b(t)$ and $Q(I)$ by

$$b^*(t) = \frac{QE(0)}{QE(b(t))} b(t) \tag{3.34}$$

$$Q^*(I) = \frac{QE(0)}{QE(Q(I))} Q(I). \tag{3.35}$$

Note that both the forward and reverse solutions involve solving an implicit relation. For the forward problem, $b(t)$ is given. It is simple to calculate $b^*(t)$ and $Q^*(I)$, but solving for $Q(I)$

involves solving (3.35), an implicit relationship. For the reverse problem, $Q(I)$ is known, from which $Q^*(I)$ and $b^*(t)$ are easily derived. Finding $b(t)$ requires solving (3.34), another implicit relationship. The implicit relations in the forward and reverse problems should be easy to solve by iteration, since $QE(Q)$ is only a weak function of Q.

Note that charge sense nonlinearity is inherently compressive. It is not possible for the algorithm to undo this compression, so it is theoretically possible for some compression curves to be unobtainable via the algorithm. This is not a case of practical interest, since the compression obtained solely from the charge sense nonlinearity is modest compared to the desired compression. As the photodiode changes voltage from 3.7 V to 2.0 V, the charge sense capacitance increases from 9.87 fF to 10.9 fF. The sensitivity of output voltage to illumination decreases approximately 9% due solely to increasing charge sense capacitance. In summary, nonlinear charge sense capacitance can be easily accounted for in the solutions of both forward and reverse problems. Any practical compression curve is still obtainable by a slight modification of the barrier curve.

Nonlinear quantum efficiency can be handled in the same way. Quantum efficiency is a weak function of voltage, since the *collection* region, but not the photogeneration region, shrinks as voltage decreases. This effect is modest, however, compared to the nonlinearity of the charge sense capacitance.

3.4.2 Clock Feedthrough

The barrier device is a MOS transistor with overlap capacitance $C_{ol,gd}$ between its gate and drain, and overlap capacitance $C_{ol,gs}$ between its gate and source. The source node is the high impedance charge sense node, so a change in b(t) can cause a significant change in the charge sense node voltage via feedthrough from the gate. The effect depends on whether or not the charge integration curve intersects the barrier curve.

Case 1

In Case 1, the charge integration curve does not intersect the barrier curve. The equivalent circuit at all times during the integration period is shown in Fig. 3-10. The sense capacitance C_{sense} includes the feedthrough capacitance $C_{ol,gs}$. The sense node voltage at the end of the integration period is the sum of the photocurrent and feedthrough components.

$$V_{sense} = IT + \frac{C_{ol,gs}}{C_{sense}} \Delta V_g \tag{3.36}$$

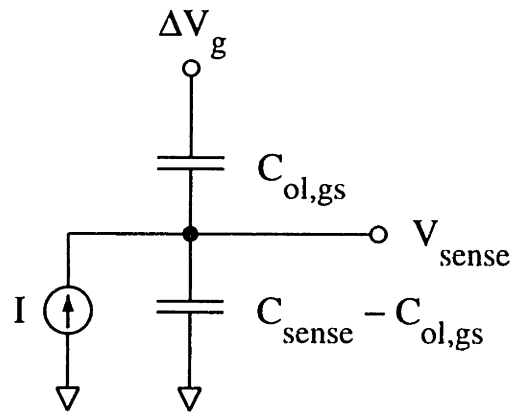


Figure 3-10: Equivalent circuit for analyzing feedthrough capacitance. V_g is the voltage on the barrier gate.

$$= IT + n\frac{C_{ol,gs}}{C_{sense}}(Q_{max} - b_{min}), \tag{3.37}$$

where V_g is the gate voltage and b_{min} is the minimum of $b(t)$. Subthreshold conduction parameter n and sense node capacitance C_{sense} depend on the sense node voltage, but are assumed constant for simplicity. To first order, the feedthrough produces a constant offset for low illumination.

Case 2

In Case 2, the charge integration curve intersects the barrier curve. While the charge integration curve is barrier limited, the feedthrough capacitance changes the effective current to

$$i_{eff} = I + nb'(t)\frac{C_{ol,gs}}{C_{sense}}. \tag{3.38}$$

At the departure time $t_{d,eff}$, the effective illumination is

$$I_{eff} = \frac{I}{1 - nC_{ol,gs}/C_{sense}}. \tag{3.39}$$

The increase in barrier-limited charge Q_{bl} due to feedthrough is

$$\Delta Q_{bl} = (Q(I_{eff}) - Q'(I_{eff})I_{eff}) - (Q(I) - Q'(I)I). \tag{3.40}$$

The change in freely accumulated charge Q_{free} due to feedthrough has two components. One component accounts for the change in gate voltage during the free integration period, and

the other accounts for the time difference of the free integration period.

$$\Delta Q_{free} = n\frac{C_{ol,gs}}{C_{sense}}(Q_{max} - b(t_{d,eff}) + I(t_d - t_{d,eff}) \quad (3.41)$$

$$= n\frac{C_{ol,gs}}{C_{sense}}(Q_{max} - b(t_{d,eff}) + I(Q'(I_{eff}) - Q'(I)) \quad (3.42)$$

The total change in charge is

$$\Delta Q_{tot} = n\frac{C_{ol,gs}}{C_{sense}}(Q_{max} - b(t_{d,eff}) + Q'(I_{eff})(I - I_{eff}) + Q(I_{eff}) - Q(I) \quad (3.43)$$

$$\approx n\frac{C_{ol,gs}}{C_{sense}}(Q_{max} - (Q(I_{eff}) - Q'(I_{eff})I_{eff}), \quad (3.44)$$

where the last three terms approximately cancel if $I \approx I_{eff}$. Given a $b(t)$, the ideal compression curve $Q(I)$ can be determined. The actual compression curve is found by rescaling the I (x) axis and determining the offset from the ideal curve at each value of I. In the reverse problem, the desired compression curve is given. The ideal compression curve $Q(I)$ must be found from this by iteration, and then the appropriate $b(t)$ can be calculated.

3.4.3 Optical Crosstalk

The effective pixel photocurrent is a weighted sum of the photocurrents at the pixel and its immediate neighbors. For uniform illumination, the effective photocurrent is

$$I_{ct,eff} = (1 + \delta)I, \quad (3.45)$$

where δ is typically around 0.01 or 0.02, and I is the photocurrent from the light entering a single pixel.

The situation is very different when the pixel has a storage region. Crosstalk charge entering the storage diffusion stays there, adding directly to the sensed charge. The total charge is then

$$Q_{tot}(I) = Q(I_{ct,eff}) + cIT, \quad (3.46)$$

where c depends on the physical layout of the pixel and process details. It will typically on the order of 1%. The compression curve with and without a storage region is shown in Fig. 3-11. The crosstalk severely limits the achievable compression. In Fig. 3-11, the maximum photocurrent with and without crosstalk are shown as $I_{max,ct}$ and I_{max}, respectively. For example, if $c = 0.01$, then the maximum improvement in dynamic range is a factor of 100, and even this would impose only one choice for the compression curve. This was the main reason for choosing a pixel without a storage region.

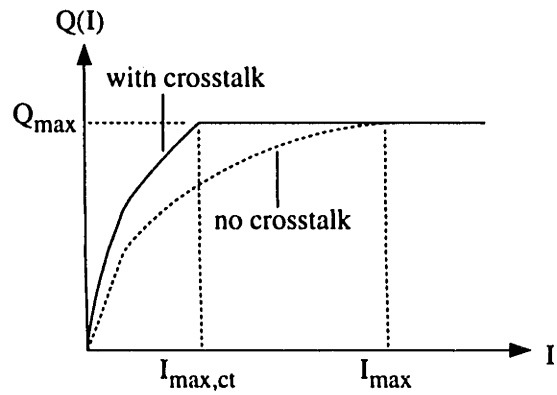


Figure 3-11: Compression curve shift with optical crosstalk.

3.4.4 Subthreshold Conduction

The barrier does not act as an infinitely sharp "knife edge." Because of subthreshold conduction, current can pass through the barrier when the charge integration curve is below the barrier. The subthreshold current is given by [50] as

$$I_D = \frac{W}{L}\mu\phi_t(-Q'_{I,source})(1 - e^{-V_{DS}/\phi_t}), \tag{3.47}$$

where

$$Q'_{I,source} = -\frac{\gamma C'_{ox}}{2\sqrt{\psi_{sa}(V_{GB})}}\phi_t e^{[\psi_{sa}(V_{GB})-2\phi_F]/\phi_t} e^{-V_{SB}/\phi_t} \tag{3.48}$$

$$\phi_t = \frac{kT}{q} \tag{3.49}$$

$$\gamma = \text{body-effect coefficient} \tag{3.50}$$

$$C'_{ox} = \text{gate capacitance per unit area} \tag{3.51}$$

$$\psi_{sa}(V_{GB}) = \text{surface potential at gate-bulk voltage } V_{GB} \tag{3.52}$$

$$\phi_F = \text{Fermi potential} \tag{3.53}$$

$$V_{SB} = \text{source-bulk voltage} \tag{3.54}$$

Assuming that the drain voltage is several hundred millivolts above the source voltage, this simplifies to

$$I_D = \frac{W}{L}\mu\phi_t\frac{\gamma C'_{ox}}{2\sqrt{\psi_{sa}(V_{GB})}}\phi_t e^{[\psi_{sa}(V_{GB})-2\phi_F]/\phi_t} e^{-V_{SB}/\phi_t} \tag{3.55}$$

$$\approx I_s e^{V_{GB}/n\phi_t} e^{-V_{SB}/\phi_t}. \tag{3.56}$$

$$\approx I_s e^{-(b(t)-q(t))/C_{pd}\phi_t}, \tag{3.57}$$

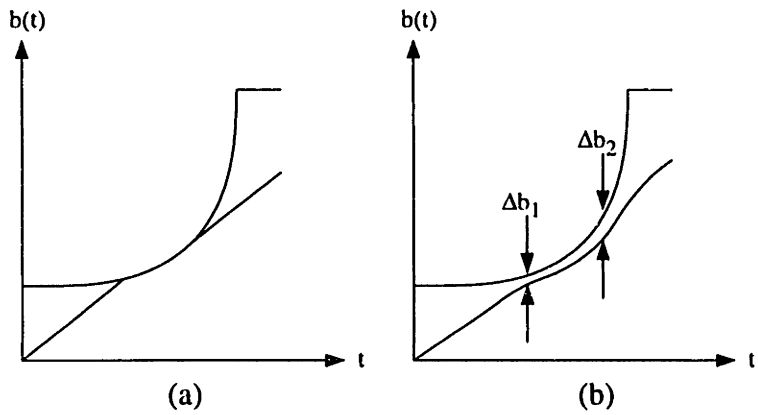


Figure 3-12: (a) Ideal charge integration curve (b) Charge integration curve with subthreshold conduction over barrier.

where the photodiode capacitance C_{pd} is approximately constant and $q(t)$ is the charge integration curve. The I_s term has a gate voltage dependency, which is weak compared to the exponential term. The gate voltage uniquely determines the channel potential, which sets the barrier height. The current then has an exponential dependence on the barrier-photodiode potential difference.

Subthreshold conduction has no effect on integrated charge for low illumination because the charge integration curve is far below the barrier, and the subthreshold current is negligible. For high illumination, the charge curve never meets the barrier curve, but instead is spaced from it by constantly varying amount. Figure 3-12 compares the charge curves for the ideal case and the subthreshold conduction case. With subthreshold current, the separation between barrier-limited charge integration and free integration becomes indistinct. During the time roughly corresponding to barrier-limited integration, the separation between the barrier and charge curves is related to the slope of $b(t)$, which is approximately equal to the slope of the charge integration curve. When the slope is small, the separation Δb_1 is small. When the slope is large, the separation Δb_2 is large.

The photocurrent splits between the subthreshold current and photodiode charging current, giving

$$I = I_s e^{-(b(t)-q(t))/C_{pd}\phi_t} + \frac{dq(t)}{dt}. \tag{3.58}$$

There is no obvious closed-form solution to this nonlinear differential equation. The equation can be solved numerically for a given barrier function $b(t)$, for example, by using a circuit simulation program such as HSPICE. The results of one such simulation for the exponential barrier function of Eqn. 3.7. Subthreshold conduction has only a modest effect, which suggests

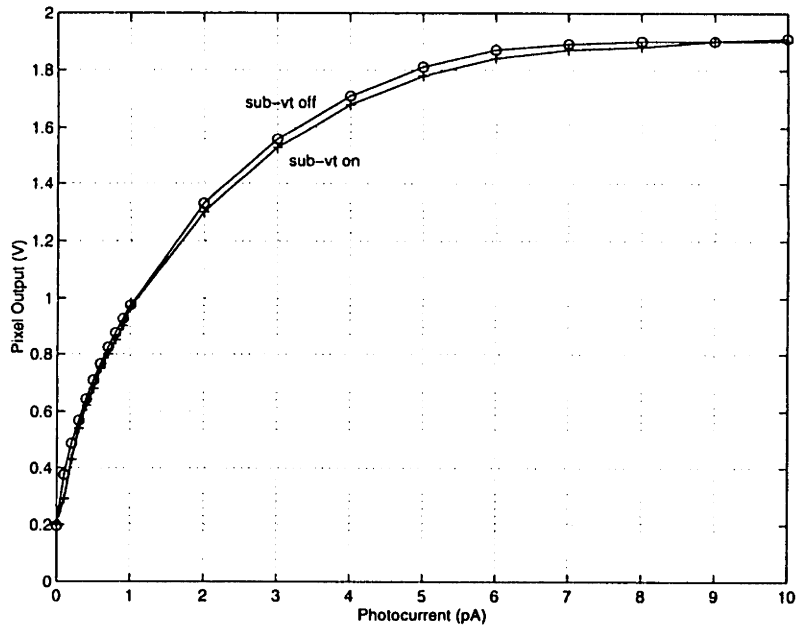


Figure 3-13: Simulated effect of subthreshold conduction on charge compression curve.

that to first order, the barrier function can be determined from the desired compression curve by ignoring subthreshold conduction. If greater accuracy is required, the barrier function must be iteratively modified until the simulated compression curve is sufficiently close to the desired compression curve.

3.5 Random Noise

Random noise sources include photonic shot noise, dark current shot noise, charge flow over the barrier, and pixel source follower thermal noise. These noise sources are independent, and can be added in quadrature to find the total random noise.

3.5.1 Photonic and Dark Shot Noise

Photonic and dark current noise is generated only during free charge integration. The associated shot noise variance is equal to the number of integrated carriers.

$$\sigma^2(N_e) = N_e \tag{3.59}$$

$$N_{e,noise,shot}(I) = \sqrt{\frac{1}{q}(I + I_{dark})(T - t_d(I))} \tag{3.60}$$

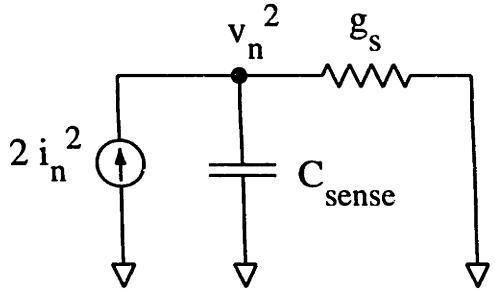


Figure 3-14: Barrier current noise model.

$$= \sqrt{\frac{1}{q}(I + I_{dark})Q'(I)}. \tag{3.61}$$

3.5.2 Barrier Current Noise

Figure 3-14 shows a noise model for the pixel when current flows over the barrier. The current through the charge spill gate has shot noise, modelled by the i_n^2 current source. The current through the barrier device also has a shot noise of the same magnitude (since the currents are the same). The voltage noise is calculated by integrating the product of the current noise spectral density and the impedance at the sense node.

$$v_n^2 = 2\int_0^\infty \frac{2qI}{|g_s + j\omega C_{sense}|^2} df \tag{3.62}$$

$$= (4qI)\frac{g_s}{C_{sense}}\frac{1}{2\pi}\frac{\pi}{2}\left(\frac{1}{g_s}\right)^2 \tag{3.63}$$

$$= \frac{qI}{g_s C_{sense}} \tag{3.64}$$

$$= \frac{kT}{C_{sense}}, \tag{3.65}$$

where g_s is the source conductance of the barrier transistor. This number is doubled by the CDS operation, since the shot noises are uncorrelated between the intervals when charge flows over the barrier. The equivalent number of noise electrons is

$$N_{e,barr}^2 = 2\frac{1}{q}\sqrt{kTC_{sense}}. \tag{3.66}$$

3.5.3 Thermal Noise

The source follower and current source devices have thermal and $1/f$ noise. The $1/f$ noise component is nearly completely removed by the CDS operation during readout. The thermal

noise of a MOS device with transconductance g_m is modelled by a shunt current source between the drain and source, having a mean square magnitude

$$i_n^2 = \frac{8}{3}\frac{kT}{g_m}\Delta f, \tag{3.67}$$

where Δf is the noise bandwidth.

The circuit and noise model of the source follower and current source are shown in Figs. 3-15(a) and (b). The noise currents are injected at the source follower output; the voltage noise is dependent on the output impedance. The output capacitance is the sum of the line capacitance C_{line} and the series combination of the photodiode capacitance C_{pd} and M1's gate-source capacitance C_{gs1}.

$$C_{out} = C_{line} + \frac{C_{pd}C_{gs1}}{C_{pd} + C_{gs1}}. \tag{3.68}$$

The output conductance is the sum of M1's back-gate transconductance g_{mb1} and the effective source conductance of M1.

$$G_{out} = g_{mb1} + g_{m1}\frac{C_{pd}}{C_{pd} + C_{gs1}}. \tag{3.69}$$

In calculating output conductance, it is important to take into account the local feedback around M1 formed by the capacitive divider C_{gs1} - C_{pd} [51].

The rms number of thermal noise electrons is

$$N_{e,thermal} = \frac{C_{sense}}{qG_{sf}}\sqrt{\frac{8}{3}\frac{kT}{G_{out}^2}(g_{m1} + g_{m2})\Delta f} \tag{3.70}$$

$$= \frac{C_{sense}}{qG_{sf}}\sqrt{\frac{8}{3}\frac{kT}{G_{out}^2}(g_{m1} + g_{m2})\frac{G_{out}}{2\pi C_{out}}\frac{\pi}{2}} \tag{3.71}$$

$$= \frac{C_{sense}}{qG_{sf}}\sqrt{\frac{2}{3}\frac{kT}{G_{out}C_{out}}(g_{m1} + g_{m2})} \tag{3.72}$$

$$= \frac{C_{sense}}{qG_{sf}}\sqrt{\frac{2}{3}\frac{kT}{(g_{mb1} + g_{m1}\frac{C_{pd}}{C_{pd}+C_{gs1}})(C_{line} + \frac{C_{pd}C_{gs1}}{C_{pd}+C_{gs1}})}(g_{m1} + g_{m2})} \tag{3.73}$$

$$\approx \frac{C_{sense}}{q}\sqrt{\frac{kT}{C_{line}}\frac{g_{m1} + g_{m2}}{g_{mb1} + g_{m1}\frac{C_{pd}}{C_{pd}+C_{gs1}}}} \tag{3.74}$$

where C_{sense} is given by

$$C_{sense} = C_{pd} + C_{gs1}(1 - G_{sf}) \tag{3.75}$$

$$= C_{pd} + C_{gs1}\frac{g_{mb1}}{g_{m1} + g_{mb1}}. \tag{3.76}$$

Since the line capacitance is much greater than the sense capacitance, thermal noise in the source follower and current source are much smaller than the barrier current noise.

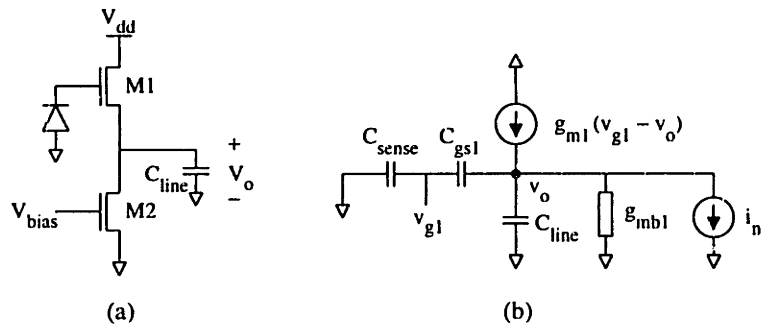


Figure 3-15: Pixel noise model for thermal noise in source follower and current source.

3.6 Fixed-Pattern Noise

Fixed-pattern noise (FPN) is caused by mismatches in the photodiode capacitance C_{pd}, optical aperture A_{opt}, pixel dark current I_d, M1's width and length W_1 and L_1, M1's effective threshold voltage $V_{te,1}$, M3's drain overlap capacitance $C_{ol,3}$, M4's effective threshold voltage $V_{te,4}$, and M4's overlap capacitance $C_{ol,4}$.

The resultant FPN was found in a previous chapter when the lateral overflow gate was constant over the integration period. This pixel differs in that most FPN terms are associated only with free charge. Table 3-1 lists the sources of FPN, their expected magnitudes, and the sensitivity of the pixel output voltage to each mismatch. The FPN with and without CDS is plotted in Fig. 3.6, assuming the compression curve of Eq. (3.7).

The sensitivity of the FPN to each source is shown qualitatively in Fig. 3-16. Mismatches in I_{bias}, $V_{t,M1}$, and $\Delta V_{t,M2}$ produce the same output offset voltage at all illuminations. Mismatches in A_{opt} and C_{sense} only affect the freely accumulated charge. There is no freely accumulated charge in the dark, so these sources do not contribute to dark FPN. At high illumination, the FPN sensitivity depends on the barrier curve; the plots in Fig. 3-16 assume that the free charge component tends towards zero as the pixel saturates. Mismatches in dark current and reset device (M2) gate-drain overlap capacitance create offset FPN which cannot be removed using CDS. The dark charge is dumped out at reset, making its contribution different during the sample and reset phases. The overlap capacitance cannot be correlated out because the reset gate voltage is different during the sample and reset phases.

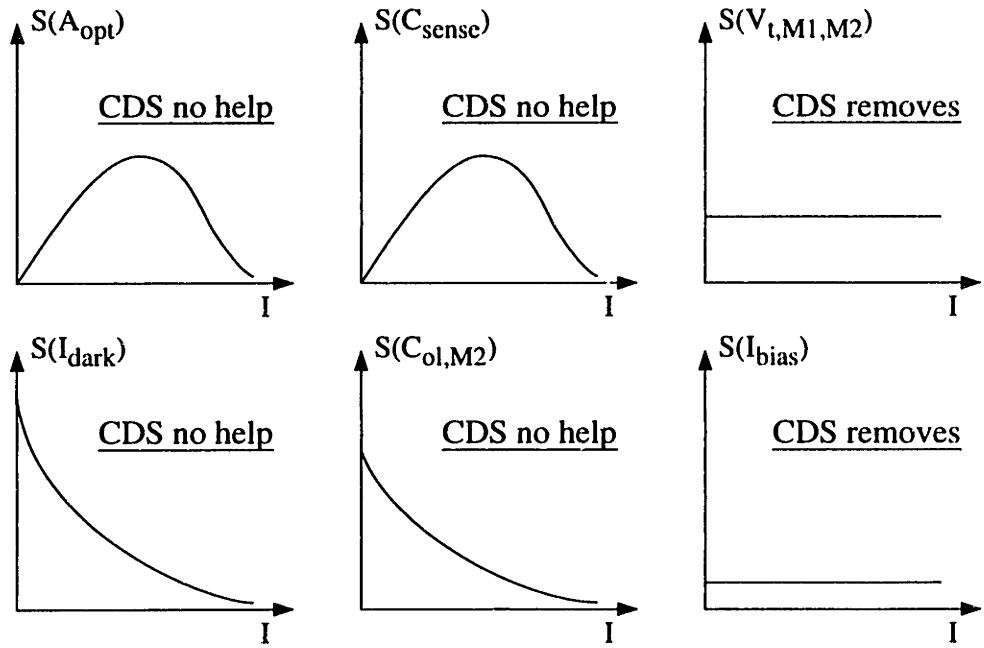
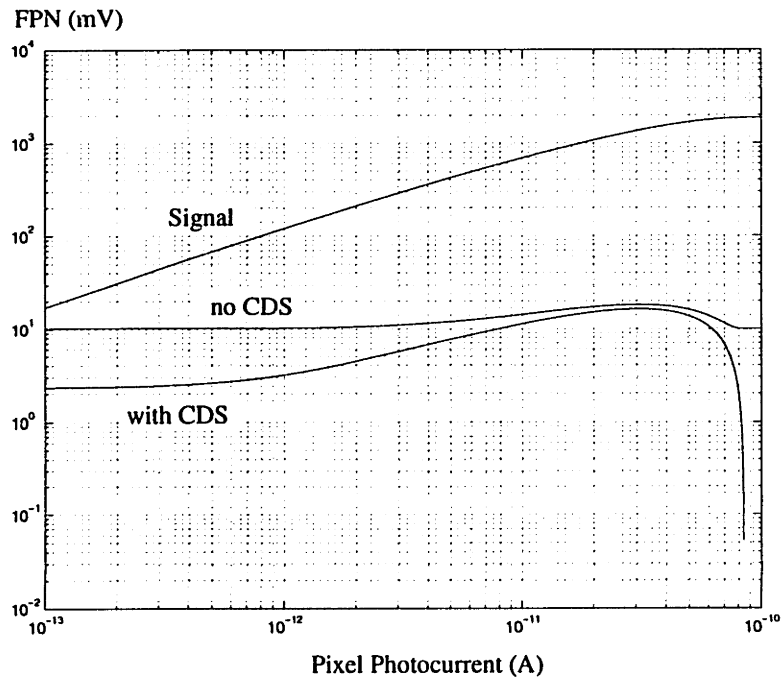


Figure 3-16: FPN sensitivities for various mismatch sources in wide dynamic range pixel.

Table 3-1: FPN sources in wide dynamic range pixel.

Parameter	Value	Sensitivity	Mismatch
C_{pd}	8.4 fF	$\frac{G_{sf}Q_{free}}{C_{sense}^2}$	2% C_{pd}
A_{opt}	237 μm^2	$-G_{sf}\frac{Q_{free}}{A_{opt}C_{sense}}\frac{I}{I+I_{dark}}$	2% A_{opt}
I_d	2.4 fA	$-G_{sf}\frac{Q_{free}/C_{sense}}{I+I_{dark}}$	10% I_d
W_1	7.1 μm	$\frac{2}{3}L_1C_{ox}(1-G_{sf})\frac{G_{sf}Q_{free}}{C_{sense}^2}+\frac{G_{sf}I_{bias}}{g_{m1}W_1}$	0.1 μm
L_1	0.8 μm	$\frac{2}{3}W_1C_{ox}(1-G_{sf})\frac{G_{sf}Q_{free}}{C_{sense}^2}-\frac{G_{sf}I_{bias}}{g_{m1}L_1}$	0.1 μm
$V_{te,1}$	1.2 V	-1	5 mV
$C_{ol,3}$	0.87 fF	$\frac{G_{sf}Q_{free}}{C_{sense}^2}$	2% $C_{ol,3}$
$V_{te,4}$	1.0 V	-1	5 mV
$C_{ol,4}$	0.87 fF	$-\frac{G_{sf}(Q_{max}-Q)}{C_{sense}^2}$	2% $C_{ol,4}$


FPN (mV)
10^4
10^3
10^2
10^1
10^0
10^{-1}
10^{-2}
Signal
no CDS
with CDS
10^{-13}
10^{-12}
10^{-11}
10^{-10}
Pixel Photocurrent (A)

Chapter 4

Experimental Chip Design

A test chip was fabricated in a 0.8 μm CMOS process with one poly layer, three metal layers, and a linear capacitor option [1]. A block diagram of the imager is shown in Fig. 4-1. The three major components of the imager are the pixel array, the barrier generation / row select block, and the upper and lower output arrays. In addition to the major blocks, the imager has a bias circuit, a two-phase clock generator, switch select and column select multiplexers, and several test structures, including an individual pixel, an individual correlated double-sampling (CDS) circuit, an individual ADC circuit, and a 10 $\times$ 10 array of photodiodes. The die photograph is shown in Fig. 4-2.

4.1 Output Circuit

The chip has 32 output pins. Each output pin is driven by one output multiplexer. Each output multiplexer combines the outputs from four ADC's, and each ADC converts the outputs of two CDS circuits. Each CDS circuit samples a single column line. A block diagram of one output multiplexer and its associated signal chain is shown in Fig. 4-3.

The op-amp is a crucial element of both the ADC and CDS circuits, so its design and performance will be explained first. The ADC and CDS circuits will be shown and their function explained. Finally, the output multiplexer and pad driver are described.

[1] Process provided by the National Semiconductor Corporation

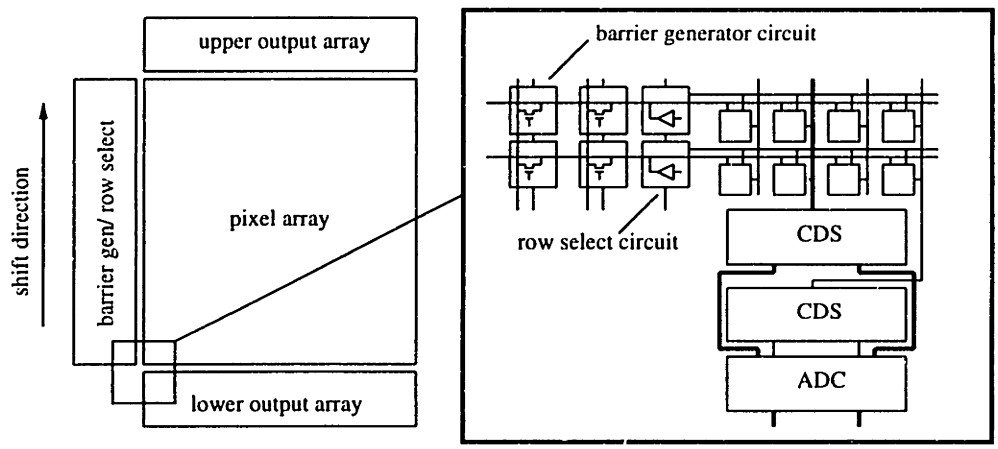


Figure 4-1: Imager block diagram. The signal path for one column is shown in bold.

4.1.1 Op-Amp

The op-amp is a folded-cascode design. Input transistors M1 and M2 form a differential pair which divides the tail current between M3 and M4. Cascodes M3 and M4 increase the effective output impedance of the input transistors M1 and M2. It will be shown later that they approximately double the DC gain. Cascodes M7 and M11 buffer the differential pair currents onto the output nodes, where they are converted to voltage by the output impedance. The circuit diagram of the core amplifier is shown in Fig. 4-4, and the bias circuit is shown in Fig. 4-5.

DC Gain

The low-frequency differential gain is found from the half-circuit consisting of M2, M4, and M5 - M8. The complete small-signal model is shown in Fig. 4-6(a). When an input v_{in} is applied to the gate of M2, the generated current is

$$i_{in} = g_{m1} v_{in}. \tag{4.1}$$

The impedance looking into the source of M4 is

$$G_{s4} = \frac{g_{m4} + g_{mb4}}{g_{o4}} \left(g_{o8} + G_p \frac{g_{m7} + g_{mb7}}{g_{o7}} \right), \tag{4.2}$$

so the fraction of i_in which passes through M4 (as opposed to flowing through the output conductance of M2) is

$$f_1 = \frac{G_4}{G_4 + g_{o2}} \tag{4.3}$$

$$= 0.92. \tag{4.4}$$

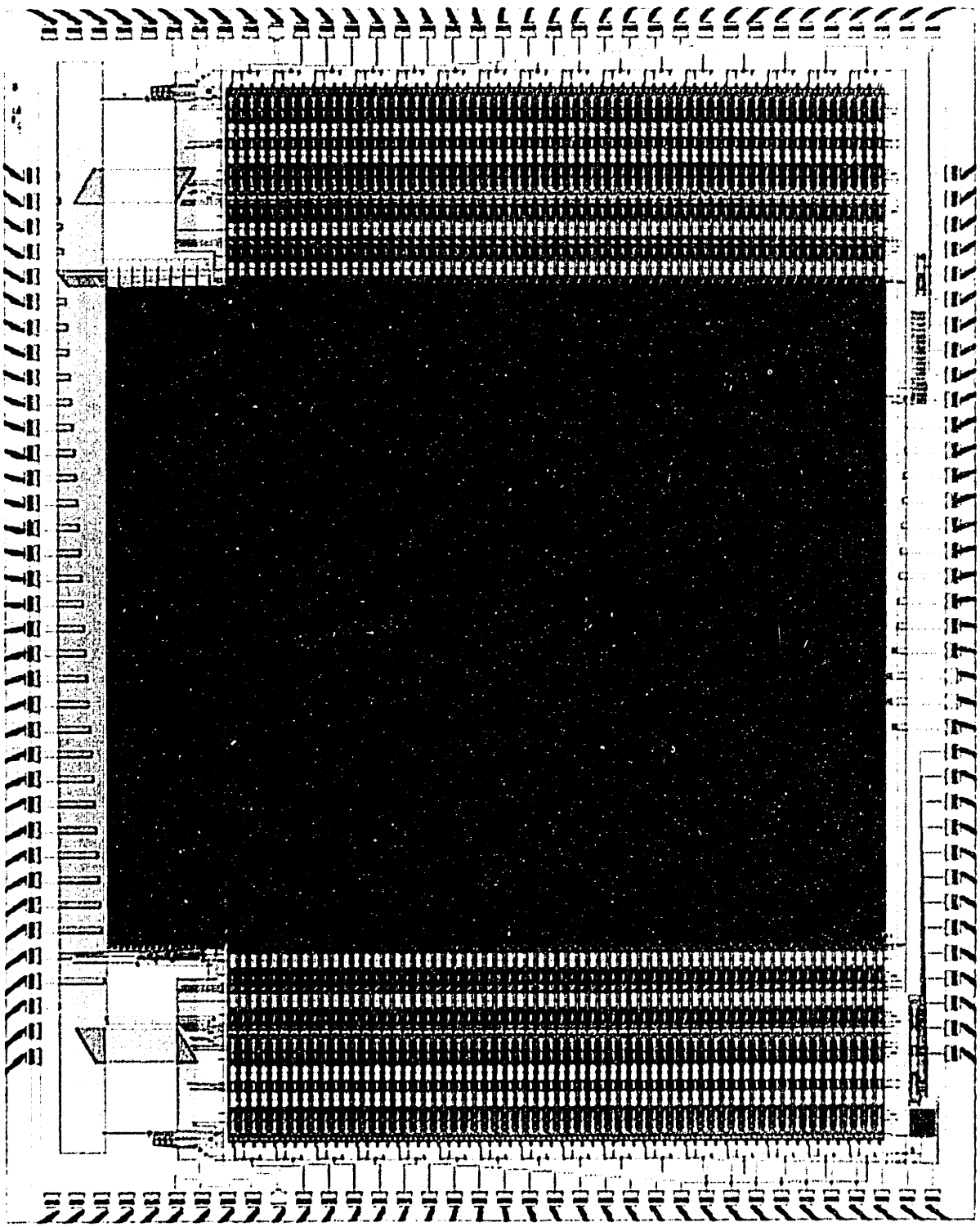


Figure 4-2: Die photograph of imager.

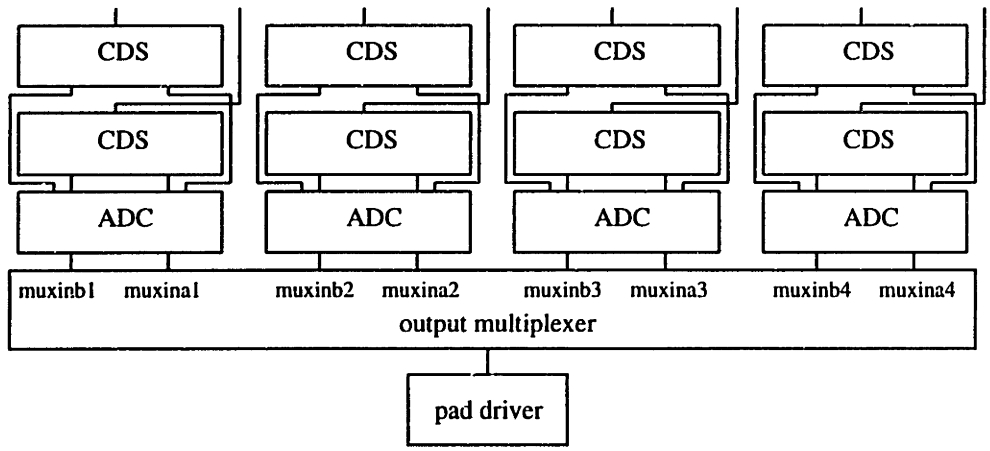


Figure 4-3: Output block diagram for one output pad.

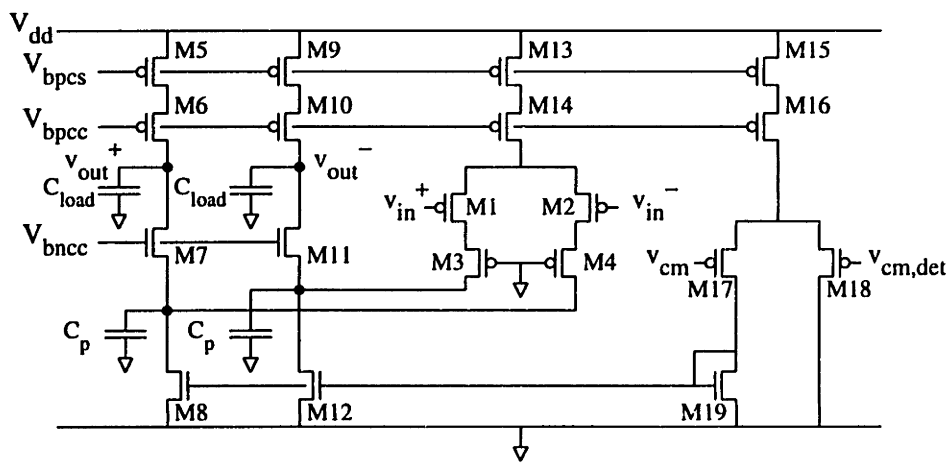


Figure 4-4: Folded cascode op-amp core.

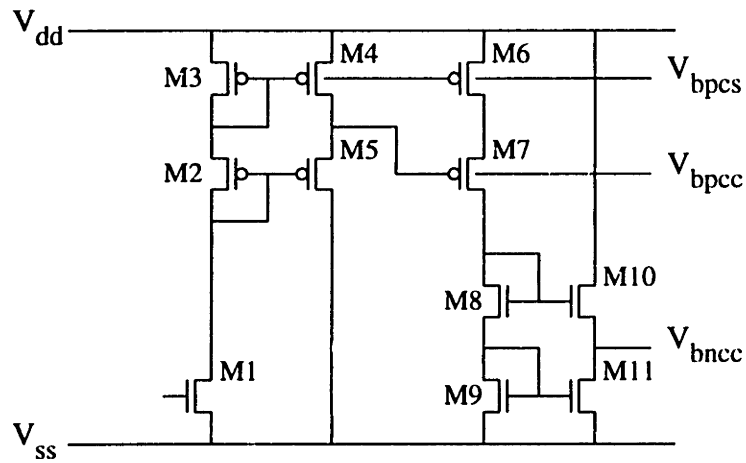


Figure 4-5: Bias circuit for folded cascode op-amp.

where f_1's value is calculated from the simulated operating point conditions.

M4's drain current splits between g_{o8} and M7. The fraction of M4's drain current which goes to M7 is

$$f_2 = \frac{G_p \frac{g_{m7}+g_{mb7}}{g_{o7}}}{G_p \frac{g_{m7}+g_{mb7}}{g_{o7}} + g_{o8}} \tag{4.5}$$

$$= 0.64. \tag{4.6}$$

The dc gain is

$$\frac{v_{out}}{v_{in}} = g_{m1} \cdot f_1 \cdot f_2 / G_p \tag{4.7}$$

$$= (204\ \mu S)(0.92)(0.64)/(5.16\ nS) \tag{4.8}$$

$$= 23,000, \tag{4.9}$$

which agrees well with the simulated dc gain of 20,200.

Cascode device M4 increases dc gain by increasing the fraction of i_{in} which reaches the source of M7. Because M2 is a minimum-length device, its output conductance is quite high. If M4 were removed, the conductances at the source of M7 would be g_{o2} (6.8 μS), g_{o8} (1.1 μS), and the conductance looking into M7's source (1.9 μS). Most of the injected current i_{in} would pass through M2's source. M4 effectively forces the current to split between M8 and M7, instead of between M8, M7, and M2.

M4 does not significantly increase op-amp area and does not require an additional bias voltage. It also does not significantly affect the input common-mode swing. To keep the input

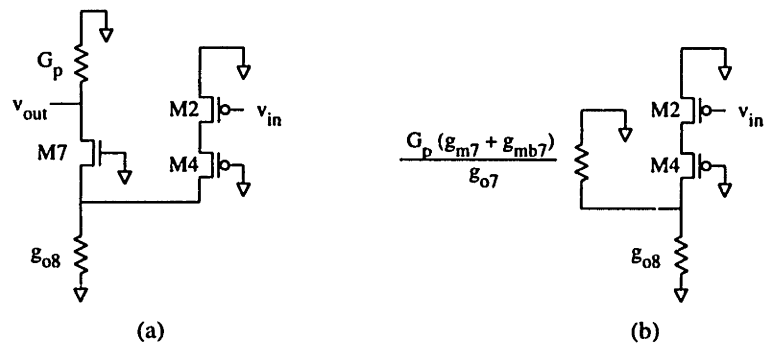


Figure 4-6: Small-signal model for evaluation of op-amp dc gain. (a) Half-circuit model. G_p is the PMOS cascode conductance. (b) Same as (a), except M7 and PMOS cascode are replaced by effective conductance.

pair saturated, the inputs cannot go below the gate overdrive voltage of M4. Under the bias conditions used, this is about 200 mV. Simulation shows that adding M4 increases DC open-loop gain from 8,380 to 20,200.

A side benefit of M4 is that it slightly reduces the input capacitance. Input capacitance is the sum of the gate-source capacitance and the Miller-multiplied gate-drain capacitance. Since M4 provides a low impedance node for M1's drain current, the Miller multiplication factor is greatly reduced. Simulation shows that adding M4 decreases parasitic capacitance at each input node from 89 fF to 78.7 fF, a minor improvement.

Differential Mode Frequency Response

The transfer function has two poles and one zero. The dominant pole is formed by the load capacitance, and is located at

$$p_1 = -\frac{G_o}{C_{load}}. \tag{4.10}$$

The nondominant pole is formed by parasitic capacitance at M7's source, and is located at

$$p_2 = -\frac{1}{C_p}\left(g_{m7} + g_{mb7} + g_{o8} + \frac{g_{o4}g_{o2}}{g_{m4} + g_{mb4}}\right) \tag{4.11}$$

$$\approx -\frac{1}{C_p}\left(g_{m7} + g_{mb7} + g_{o8}\right). \tag{4.12}$$

A zero at $z_1 = g_{m2}/C_{gd2}$ is formed by the gate-drain overlap capacitance of the input device M2. At z_1, the transconductance current $g_{m2}v_{in}^-$ is cancelled by the displacement current through C_{gd2}, so that no current goes to the output. A Bode magnitude plot of the differential mode

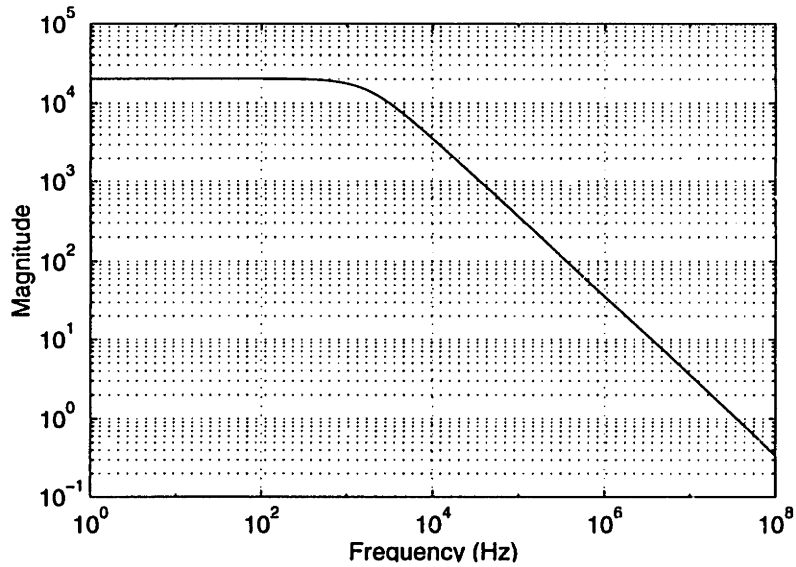


Figure 4-7: Simulated op-amp differential mode response.

response is shown in Fig. 4-7. An additional 500 fF capacitor has been added to each output in this simulation to account for loading.

Common-Mode Feedback

Any fully differential op-amp requires common-mode feedback to maintain operation in the high gain region. The common-mode feedback consists of a common-mode detector and feedback network. The common-mode detector is shown in Fig. 4-8. If **phi5** is low and $V_{cm,det}$ is initially midway between V_{out}^{+} and V_{out}^{-}, then any changes in the output common mode will be tracked by $V_{cm,det}$.

At startup, or in the presence of leakage current, the voltage at node A can differ from the output common mode. Capacitors C3 - C4 and switch transistors M1 - M6 act as a switched-capacitor resistor which restores $V_{cm,det}$ to the output common mode. The clocking frequency must be high enough to prevent leakage current from significantly changing the detected common mode.

The common-mode detector is used in a feedback loop which includes M17 - M19, M8 and M12 in Fig. 4-4. As the output common mode increases, $V_{cm,det}$ increases. More of the tail current is steered through M17 and M19, increasing the current through M8 and M12. This, in turn, pulls down the outputs, restoring the output common mode.

The gain and frequency response of the common mode loop are important since that loop

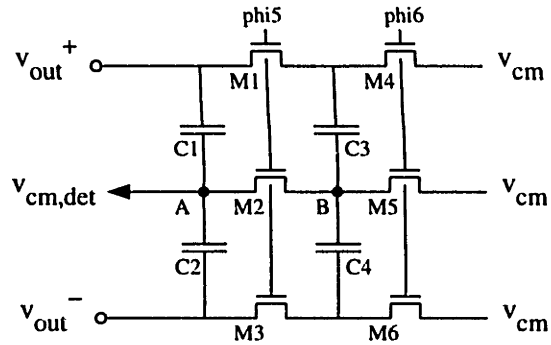


Figure 4-8: Common-mode detector.

must be stable, but cannot be much slower than the differential response. The dc gain of the common mode loop is

$$A_{cm,dc} = \frac{2C1}{2C1 + C_{g18}} \frac{g_{m17}}{g_{m19}} \frac{g_{m12}}{g_{m1}} A_{dm,dc}, \tag{4.13}$$

where $A_{dm,dc}$ is the differential mode gain at dc and C_{g18} is the total capacitance at the gate of M18.

The common mode feedback has the same poles as the differential response, plus one more due to the transconductance and capacitance at the gates of M8 and M12. The extra pole tends to occur before the nondominant pole of the differential response, which makes it necessary to limit the gain of the common mode loop. Extra parasitic capacitance was intentionally added at the gate of M18 for this purpose.

$$\begin{aligned} p_1 &= -\frac{G_o}{C_{load}} \\ p_2 &= -\frac{g_{m19}}{C_{g19} + C_{g8} + C_{g12}} \\ p_3 &= -\frac{1}{C_{p1}} \left(g_{m7} + g_{mb7} + g_{o8} + \frac{g_{o4} g_{o2}}{g_{m4} + g_{mb4}} \right). \end{aligned}$$

A plot of the simulated common mode response is shown in Fig. 4-9.

Common Mode Rejection

In the ADC and CDS, the op-amp differential offset is removed during the reset/sample phase by putting the op-amp in unity-gain feedback. The input common mode during reset is, of course, identical to the output common mode. The input common mode may shift during subsequent phases due to charge injection or variation in output common mode in different op-amps. The op-amp used in the CDS circuit exhibits a particularly large shift in input common

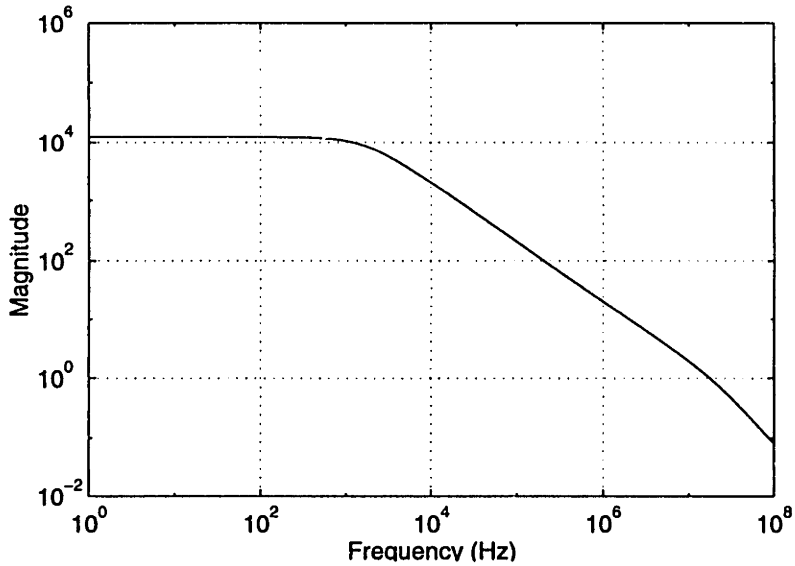


Figure 4-9: Simulated op-amp common-mode response.

mode between its **sample1** and **sample2** phases. Due to device mismatch, the change in input common mode results in a change in output differential mode, which appears as an error voltage.

The small-signal model for the input stage is shown in Fig. 4-10. The tail current source is assumed to be ideal, *i.e.*, have zero output conductance. Each input pair device parameter is assumed to consist of a common-mode component and a differential-mode component.

$$g_{m1} = g_{m,i} - \frac{1}{2}\Delta g_{m,i} \tag{4.14}$$

$$g_{m2} = g_{m,i} + \frac{1}{2}\Delta g_{m,i} \tag{4.15}$$

$$g_{o1} = g_{o,i} + \frac{1}{2}\Delta g_{o,i} \tag{4.16}$$

$$g_{o2} = g_{o,i} - \frac{1}{2}\Delta g_{o,i}, \tag{4.17}$$

where $g_{m,i}$ is the nominal transconductance of input devices M1 and M2, and $g_{o,i}$ is the nominal drain-source conductance of input devices M1 and M2.

A common-mode signal $v_{i,cm}$ and a differential-mode signal $v_{i,dm}$ are simultaneously applied to the inputs so that the net output voltage does not change. In order for this to be true, the voltages at the gates of M7 and M11 in Fig. 4-4 must stay constant despite input common-mode voltage shift. It must also be true that the sources of M3 and M4 have the same voltage in order to sink the same current; however, that common voltage changes significantly with input

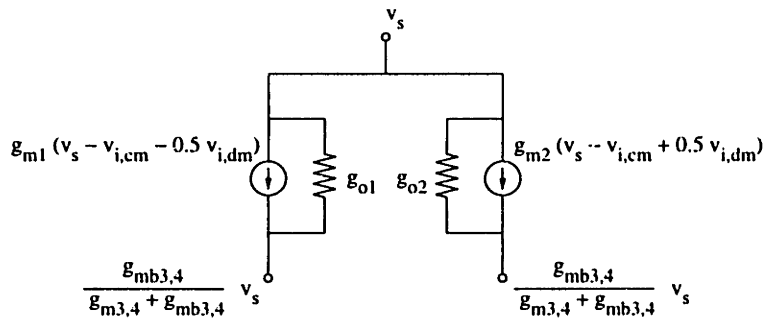


Figure 4-10: Small-signal model for input pair mismatch.

common-mode level. The drain currents of M3 and M4 cannot change with input common-mode level (ideal tail current source), but their bulk voltage changes with input common-mode level, since they are in the same well as M1 and M2. Their source voltage must also change so that the incremental currents injected by the $g_{m3,4}$ and $g_{mb3,4}$ sources cancel. The change in source voltage given a change v_s in the bulk potential is

$$v_{s3,4} = v_{d1,2} = \frac{g_{mb3,4}}{g_{m3,4} + g_{mb3,4}} v_s, \tag{4.18}$$

where $v_{s3,4}$ is the source voltage of M3 and M4, equal to the drain voltage $v_{d1,2}$ of M1 and M2.

The sum of the incremental currents in the M1 branch and M2 branch must be zero.

$$\begin{aligned} 0 &= (g_{m1} + g_{m2})(v_s - v_{i,cm}) + \frac{1}{2}(g_{m2}) - g_{m1})v_{i,dm} \\ &\quad + (g_{o1} + g_{o2})\left(v_s - v_s \frac{g_{mb3,4}}{g_{m3,4} + g_{mb3,4}}\right) \qquad (4.19) \\ &= 2g_{m,i}(v_s - v_{i,cm}) + \frac{1}{2}\Delta g_{m,i} v_{i,dm} + 2g_{o,i} v_s \frac{g_{m3,4}}{g_{m3,4} + g_{mb3,4}}. \qquad (4.20) \end{aligned}$$

It follows that the source voltage of M1 and M2 is

$$\frac{v_s}{v_{i,cm}} = \frac{g_{m,i}}{g_{m,i} + g_{o,i}\frac{g_{m3,4}}{g_{m3,4}+g_{mb3,4}}}. \tag{4.21}$$

The difference in current between the M1 and M2 branches must be zero, because $v_{i,cm}$ and $v_{i,dm}$ are chosen to make it so.

$$\begin{aligned} &g_{m1}(v_s - v_{i,cm} - \frac{1}{2}v_{i,dm}) + g_{o1} v_s \left(1 - \frac{g_{mb3,4}}{g_{m3,4} + g_{mb3,4}}\right) = \\ &g_{m2}(v_s - v_{i,cm} + \frac{1}{2}v_{i,dm}) + g_{o2} v_s \left(1 - \frac{g_{mb3,4}}{g_{m3,4} + g_{mb3,4}}\right). \qquad (4.22) \end{aligned}$$

The common-mode rejection ratio (CMRR) is defined as the ratio of differential-mode gain to common-mode gain. The common-mode gain is not the common-mode loop gain derived in Eq. 4.13, but the ratio of a common-mode input to the differential signal produced at the op-amp outputs. After some algebra, CMRR is found to be

$$\text{CMRR} = \frac{v_{i,cm}}{v_{i,dm}} = \frac{1 + \frac{g_{m,i}}{g_{o,i}}(1 + \frac{g_{mb3,4}}{g_{m3,4}})}{\frac{\Delta g_{o,i}}{g_{o,i}} + \frac{\Delta g_{m,i}}{g_{m,i}}}. \tag{4.23}$$

Input Offset Voltage

Threshold voltage and geometry mismatches create input offset voltage, which is defined as the input differential voltage which must be applied to produce zero output voltage. The systematic offset voltage is zero due to the fully differential structure of the op-amp. The random offset is due mainly to random mismatches in device pairs M1 and M2, M8 and M12, and M5 and M9.

$$\begin{aligned} \Delta v_{os} &= \Delta V_{t,M1-M2} + \frac{I_{M1}}{g_{m1}}\left(\frac{\Delta(W/L)}{W/L}\right)_{M1,2} \\ &+ \frac{g_{m8}}{g_{m1}}\left(\Delta V_{t,M8,12} + \frac{I_{M8}}{g_{m8}}\left(\frac{\Delta(W/L)}{W/L}\right)_{M8,12}\right) \\ &+ \frac{g_{m5}}{g_{m1}}\left(\Delta V_{t,M5,9} + \frac{I_{M5}}{g_{m5}}\left(\frac{\Delta(W/L)}{W/L}\right)_{M5,9}\right) \end{aligned} \tag{4.24}$$

The mismatches are independent Gaussian random variables, and can therefore be added in quadrature.

Random Noise

Thermal and $1/f$ noise in the transistors can be referred back to the input as an equivalent input noise for the op-amp. Thermal noise is assumed to be the dominant noise source, since $1/f$ noise is cancelled, to first order, in the same way that op-amp input offset voltage is cancelled.

The important thermal noise sources are M1 and M2, M8 and M12, and M5 and M9. Thermal noise in a MOSFET can be modeled as a shunt current source from drain to source.

$$\overline{i_d^2} = \frac{8}{3}kTg_m\Delta f. \tag{4.25}$$

Since the bandwidth for both the current source devices and the input devices is set by the capacitance and conductance at the outputs, the squared noise currents can be directly added. For the M2 branch of the op-amp,

$$\overline{i_n^2} = 2\frac{8}{3}kT(g_{m2} + g_{m8} + g_{m5})\Delta f, \tag{4.26}$$

Table 4-1: Summary of Operational Amplifier Design [1]

Differential DC Gain	20,200
Differential Unity Gain Bandwidth [2]	35.3 MHz
Differential Phase Margin [2]	77°
Common-Mode DC Gain	12,300
Common-Mode Unity Gain Bandwidth [2]	17.6 MHz
Common-Mode Phase Margin [2]	47°
CMRR [3]	65 dB
Power Dissipation	855 μW
Input Offset Voltage [4]	18 mV
Input Random Noise	17.6 nV/$\sqrt{Hz}$

[1] Simulated parameters.
[2] Assumes 500 fF load at outputs.
[3] Assumes 1% relative mismatch in g_m and g_o.
[4] Assumes 10 mV threshold voltage mismatch and 1% relative geometry mismatch.

where $\overline{i_n^2}$ is the total noise current. Referred to the input, and taking both differential circuits into account, this represents a voltage noise spectral density

$$S_0 = \frac{16}{3}\frac{kT}{g_{m2}^2}(g_{m2} + g_{m8} + g_{m5}). \tag{4.27}$$

The actual noise is proportional to noise bandwidth, and therefore inversely proportional to the capacitive load at the output.

4.1.2 ADC Architecture

A simplified diagram of the cyclic converter is shown in Fig. 4-11. Conversion starts by sampling the input voltage A_{in}. The input range is -1 V to +1 V. The input is compared to zero to produce the MSB b_0; b_0 is 1 if the input is positive and 0 if the input is negative. The input is doubled, $-V_{ref} = -1$V is added if b_0 is 1, and $V_{ref} = 1$V is added if b_0 is 0. In either case, the analog output (residue) from the first stage is between -1 V and +1 V. The second stage performs the same operation as the first stage, taking as its input the residue of the first stage. It produces the next most significant bit b_1. The first stage then samples the residue of the second stage, producing b_3. This cycle continues until the desired number of bits is obtained. Design techniques exist which provide insensitivity to finite op-amp gain and capacitor-ratio [52], but these techniques were not used due to the large number of extra clocking phases and capacitors required.

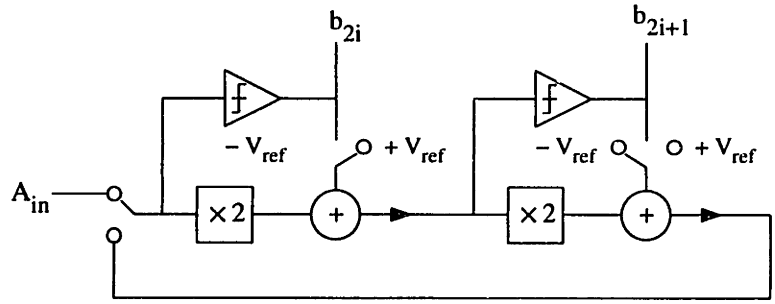


Figure 4-11: Simplified cyclic converter schematic.

A complete schematic is shown in Fig. 4-12, and a simpler schematic diagram in Fig. 4-13. The cyclic converter contains two op-amps, two comparators, eight capacitors, and several switches. A four-phase nonoverlapping clock is used, with an additional phase, ϕ**0**, used to sample the input. The circuit can be divided into a first stage consisting of op-amp #1, comparator #1, capacitors C1 - C4, and their associated switches, and a second stage consisting of op-amp #2, comparator #2, capacitors C5 - C8, and their associated switches.

Conversion starts with ϕ**0** and ϕ**1** high. Clocks with "D" suffixes are slightly delayed versions of the base clock; *e.g.*, ϕ**1D** is a delayed version of ϕ**1** and ϕ**1DD** is a delayed version of ϕ**1D**. The outputs from comparator #1, **CO1+** and **CO1-**, are both low during the first phase. Op-Amp #1 is in unity-gain feedback, and samples the op-amp input offset voltage on the top plates of C1 - C4, while the ADC input is sampled on the bottom plates. The ADC configuration during phase 1 is shown in Fig. 4-13(a).

In the second phase, ϕ**2** goes high. The feedback loop around op-amp #1 is opened, and the bottom plates of C1 - C4 are shorted together. In this phase, op-amp #1 acts as an offset-cancelled preamp for comparator #1. A finite voltage difference exists at the inputs to the open-loop op-amp, causing the outputs to diverge toward the rails. The output at A-A' is gated to the inputs of comparator #1. As ϕ**2** goes low, the comparator inputs are disconnected from the op-amp outputs. The ADC configuration during phase 2 is shown in Fig. 4-13(b).

Several events occur during the third phase. As ϕ**3** goes high, A and A' are briefly shorted together through three series switches. This assists recovery of the op-amp from its saturated state. The bottom plates of C3 and C4 are connected to the op-amp outputs. If the bottom plates of C1 and C2 were shorted together, their charge would be shifted onto C3 and C4, resulting in a doubling of the input voltage. In fact, C1 and C2 are connected to $\pm V_{ref}$. Conceptually this is equivalent to shorting the bottom plates of C1 and C2 together, followed by adding or subtracting $+V_{ref} - -V_{ref} = 2$ V.

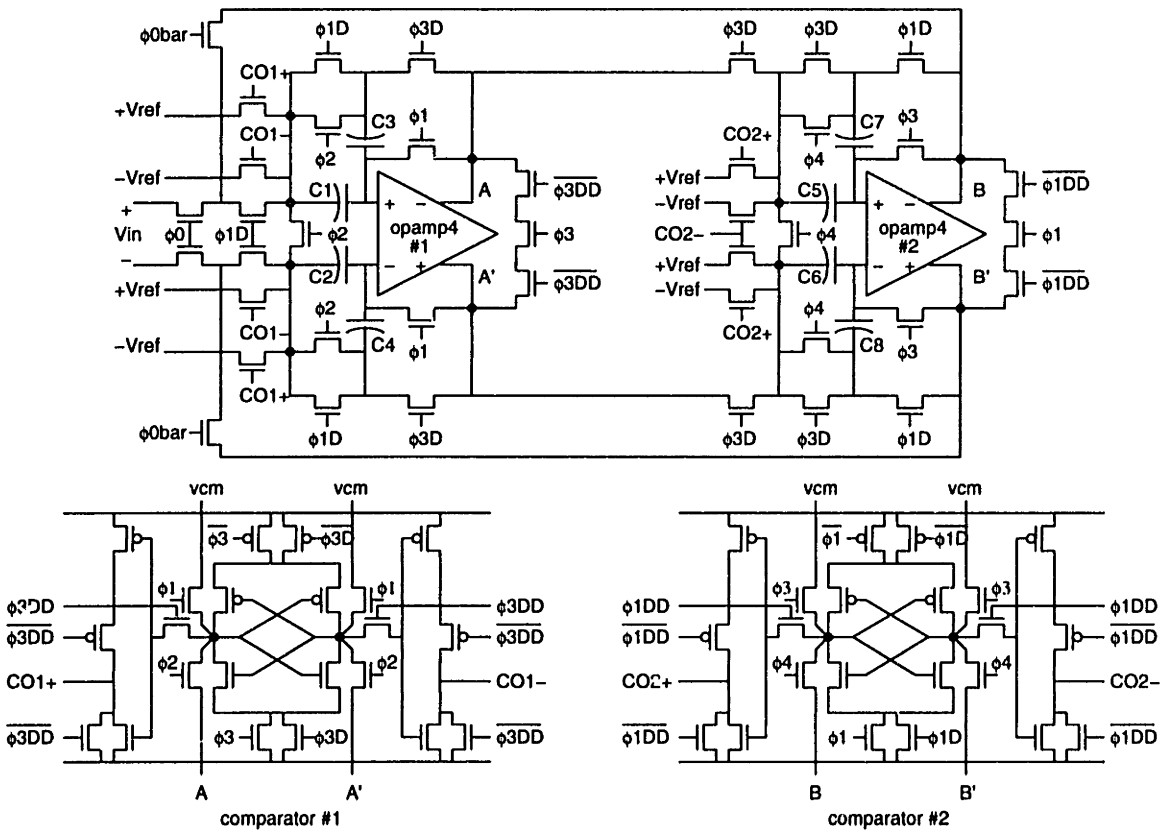


Figure 4-12: Complete cyclic ADC schematic.

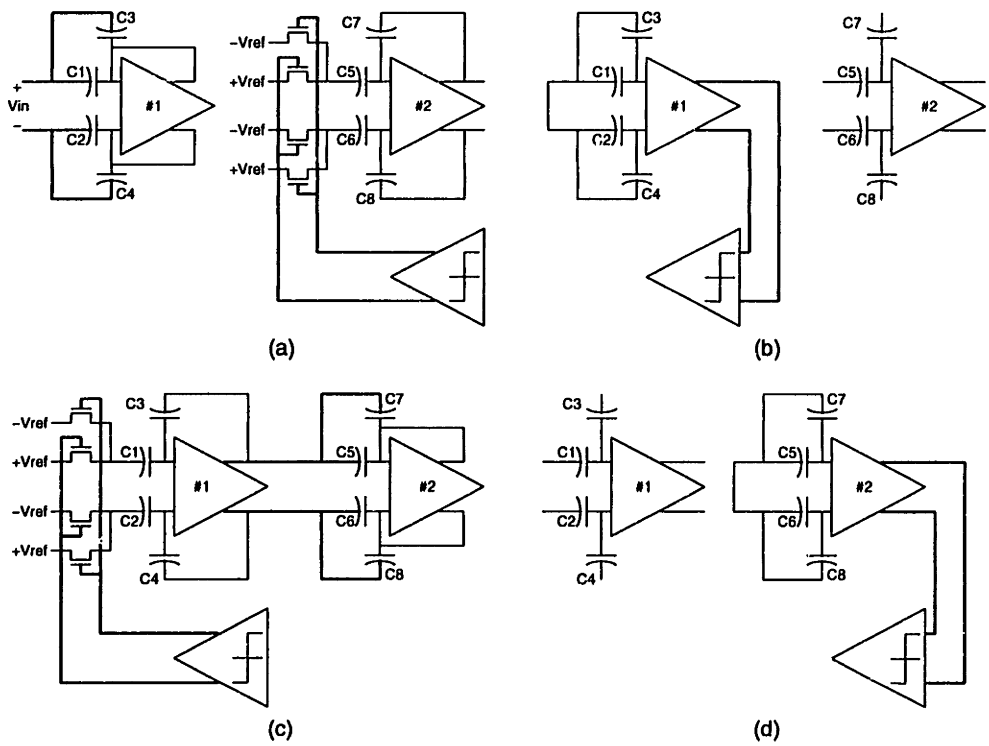


Figure 4-13: Simplified schematic for cyclic ADC. Signal path is shown in bold. (a) Phase 1 (b) Phase 2 (c) Phase 3 (d) Phase 4.

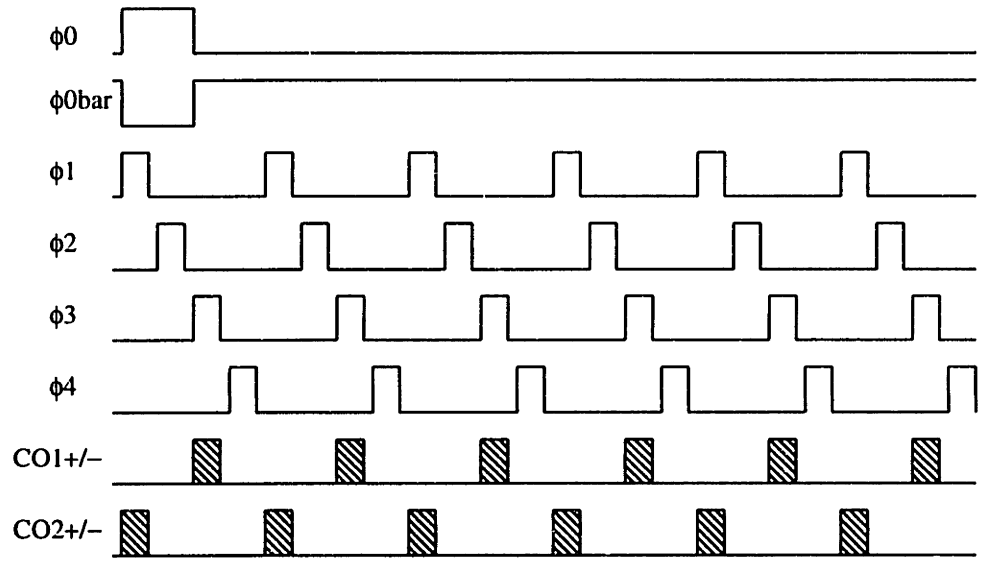


Figure 4-14: Timing for cyclic ADC.

During phase 3, the latch is connected to the power rails, causing it flip to one of two stable states. The comparator outputs, CO1+ and CO1-, are used to select whether $+V_{ref}$ is applied to C1 and $-V_{ref}$ to C2 or vice versa. The comparator output is also sent to cpouta; cpouta will be low if the input was positive and high if the input was negative. Meanwhile, the second stage is now in the same configuration that the first stage was in during the first phase. Instead of sampling the ADC input, however, the second stage samples the residue produced by the first stage. The ADC configuration during phase 3 is shown in Fig. 4-13(c).

During the fourth phase, the first stage is inactive. The second stage is in the same configuration that the first stage was in during the second phase. The sampled residue is compared to zero to calculate the second bit. This bit is sent to the output as cpoutb, and is low if the second bit is 0 and high if the second bit is 1. The ADC configuration during phase 4 is shown in Fig. 4-13(d).

The ADC returns to the first phase, and the second stage is now producing a residue voltage. The first stage samples its input, which now comes from the output of the second stage instead of the ADC input. This cycle continues until all bits have been generated. A sample timing diagram for one complete 12b conversion is shown in Fig. 4-14.

4.1.3 ADC Nonidealities

Capacitor Mismatch The algorithmic converter depends on accurate matching between capacitor pairs C1, C3 and C2, C4. The residue of one stage is

$$V_{residue} = V_{in}\left(1+\frac{1}{2}\left(\frac{C1}{C3}+\frac{C2}{C4}\right)\right) \pm V_{ref}\frac{1}{2}\left(\frac{C1}{C3}+\frac{C2}{C4}\right), \tag{4.28}$$

where $V_{ref} = V_{refp} - V_{refm}$. If the standard deviation of the capacitor value is ΔC, then the residue has variance

$$\begin{aligned}(\Delta V_{residue})^2 &= \left(\frac{\partial V_{residue}}{\partial C1}\right)^2(\Delta C1)^2 + \left(\frac{\partial V_{residue}}{\partial C2}\right)^2(\Delta C2)^2 + \left(\frac{\partial V_{residue}}{\partial C3}\right)^2(\Delta C3)^2 \\ &\quad + \left(\frac{\partial V_{residue}}{\partial C4}\right)^2(\Delta C4)^2 \qquad (4.29)\\ &= \frac{2}{C}(V_{in} \pm V_{ref})^2(\Delta C)^2 \qquad (4.30)\end{aligned}$$

The corresponding residue plot is shown in Fig. 4-15(a).

Op-Amp Input Offset Voltage The op-amp offset voltage is mostly cancelled during the sampling phase. The cancellation is incomplete due to the finite op-amp gain. The voltage at the input and output of the op-amp during the sampling phase is

$$V_{in} = \frac{A}{A+1}V_{os}. \tag{4.31}$$

The effective residual input-referred offset is

$$V_{os,res} = \frac{1}{A+1}\frac{2C+C_p}{C}V_{os}, \tag{4.32}$$

where C_p is the parasitic capacitance at the op-amp input. The corresponding residue plot is shown in Fig. 4-15(b).

Finite Op-Amp Gain Aside from incomplete cancellation of the op-amp offset voltage, finite op-amp gain causes gain error in the calculation of the residue. The op-amp places the correct charge on the capacitors during the reset phase, since the output voltage in that phase is zero, and the finite gain is irrelevant. When the residue is output from the op-amp, the finite gain requires that a small fraction of the op-amp output voltage be applied at the op-amp inputs. If the ideal output value is $V_{residue}$, the voltage at the op-amp input is $V_{residue}/A$, where A is

the op-amp gain. The input-referred error for the stage is the voltage that must be applied to C1 and C2 in order to create that op-amp input voltage. This error is

$$V_{A,err} = \frac{1}{A}\frac{2C + C_p}{C}V_{residue}. \tag{4.33}$$

The residue plot with finite op-amp gain is shown in Fig. 4-15(c).

Comparator Offset Voltage Comparator offset is important when the sampled input voltage is so small that the op-amp outputs barely move during the pre-amp phase. Moving from the stage input to the op-amp input terminals, the input voltage v_{in}, which is the residue of the previous stage, is scaled by a capacitive divider.

$$v_{in,oa} = v_{in}\frac{2C}{2C + C_p} \tag{4.34}$$

The differential output current from the op-amp is the applied input voltage $v_{in,oa}$ multiplied by the transconductance $g_{m,i}$ of the input pair devices. This current is integrated by C_{load}, which is the comparator input capacitance plus parasitics on either op-amp output node. This charging phase has duration $\delta t_{pre-amp}$. The condition for the input voltage to exceed the comparator offset is

$$|v_{in,oa}| \cdot g_{m,i} \cdot \delta t_{pre-amp} \geq C_{load}V_{comp,os} \tag{4.35}$$

$$|v_{in}| \geq V_{comp,os}\left(1 + \frac{1}{2}\frac{C_p}{C}\right)\left(\frac{C_{load}}{g_{m,i}}\right)\left(\frac{1}{\delta t_{pre-amp}}\right) \tag{4.36}$$

$$\geq V_{comp,os}\left(1 + \frac{1}{2}\frac{C_p}{C}\right)(\text{UGBW} \cdot \delta t_{pre-amp})^{-1}, \tag{4.37}$$

where UGBW is the unity-gain bandwidth of the op-amp. The residue is qualitatively identical to the plot for op-amp input offset voltage (Fig. 4-15(b)).

Charge Injection When a switch opens, it injects channel charge and displacement charge into the source and drain. The only relevant switching event is at the end of the sampling phase, when the feedback loops around the op-amp are opened. If the ADC is perfectly symmetric, the charges injected on both sides are identical, and there is no net effect. In practice, these charges differ because the switch transistors do not exactly match, their initial source/drain potentials differ due to op-amp input offset voltage, and the impedances through the sampling capacitors differ due to the differing voltages at the two inputs.

The nominal charge injected into either side is a complicated function of clock timing and source/drain impedances. It can be assumed that the fall time of the switch is small compared

to the op-amp settling time, so that the op-amp output is purely capacitive. It is also assumed that the fall time is long enough that the voltage across the switch is zero until the channel charge is completely drained. Under these assumptions, the nominal charge injected into each summing node is

$$Q_{inj} = -(WLC_{ox} + 2C_{ol})(V_{dd} - V_t - V_{cm})\frac{2C + C_p}{2C + C_p + C_{out}} - (V_t + V_{cm})C_{ol}, \tag{4.38}$$

where W and L are the width and length of the switch device, V_t is the threshold voltage with a source voltage of V_{cm}, C_{ol} is the gate-drain overlap capacitance, and C_{out} is the capacitance at the op-amp output.

The mismatch ΔQ_{inj} in the injected charge is typically on the order of 10% of Q_{inj}. The injected charge has an input voltage dependence due to the impedance mismatch between the two sides, but this is ignored. The equivalent error in V_{in} is

$$\Delta V_{in} = \frac{\Delta Q}{2C + C_p}. \tag{4.39}$$

The residue plot with charge injection error is shown in Fig. 4-15(d).

Random Noise The switches introduce kT/C noise when they open, and the op-amp itself has input-referred thermal and $1/f$ noise. The $1/f$ noise is similar to op-amp offset voltage, and is assumed to be cancelled in the same manner. Noise potentially results in errors at two points in the algorithm. First, input-referred op-amp noise causes fluctuation in the comparator input during the pre-amplification phase. Second, switch noise and op-amp thermal noise cause errors in the input sampling operation.

During the pre-amp phase, the op-amp input-referred noise is added directly to the input signal. The op-amp and capacitive load form a linear system with an impulse response as shown in Fig. 4-16. The mean square noise at the output is

$$\begin{aligned} E[v_{o,n}^2(t)] &= \int_{-\infty}^{\infty} S_{v_{i,n}}(f)|H(f)|^2 df & (4.40)\\ &= S_0 \int_{-\infty}^{\infty} |H(f)|^2 df & (4.41)\\ &= S_0 \int_{-\infty}^{\infty} h^2(t) dt & (4.42)\\ &= S_0 \left(\frac{g_m}{C_{load}}\right)^2 t_{phase}, & (4.43) \end{aligned}$$

where the input noise spectral density $S_{v_{i,n}}$ is S_0, g_m is the transconductance of the op-amp input pair, t_{phase} is the during of one phase ($\phi 1$, $\phi 2$, $\phi 3$, or $\phi 4$) of operation, and Parseval's

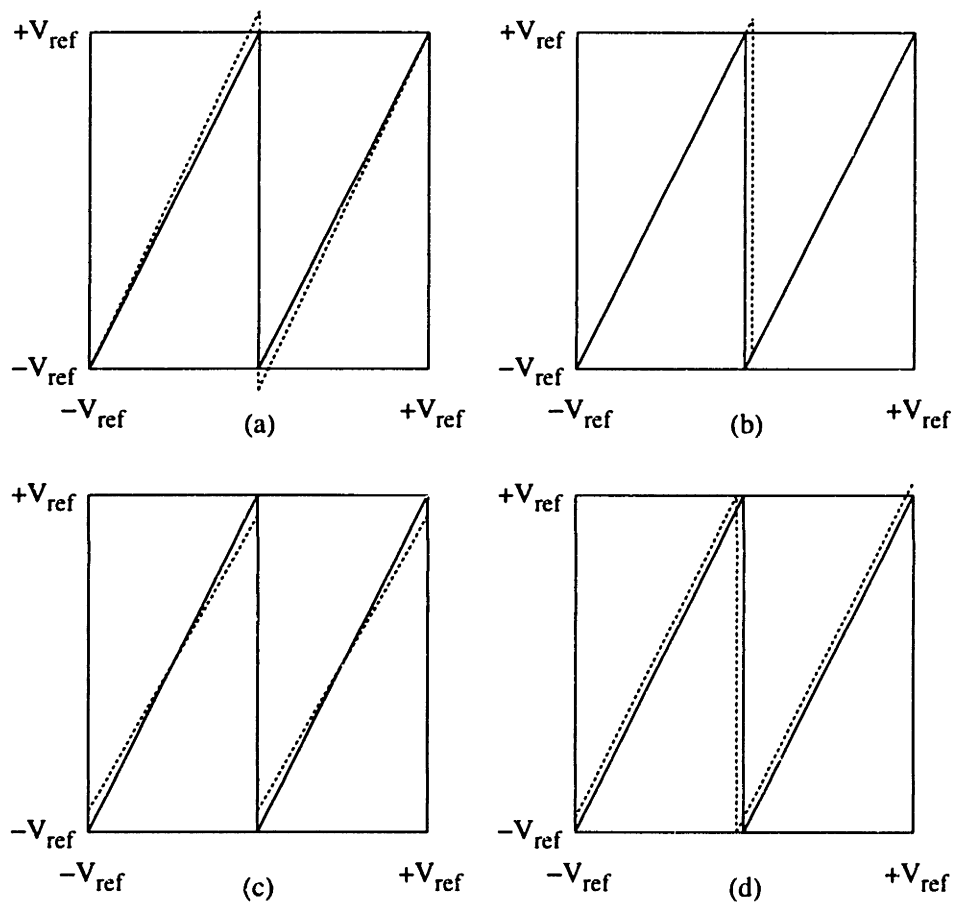


Figure 4-15: Residue plots for various nonidealities: (a) capacitor ratio mismatch (b) op-amp input offset voltage (c) finite op-amp gain (d) charge injection

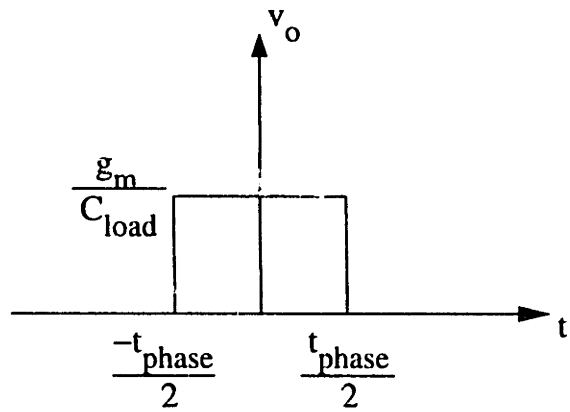


Figure 4-16: Impulse response of op-amp and capacitive load.

theorem has been applied. For the deterministic signal, the system gain is

$$\frac{v_o}{v_i} = \frac{2C}{2C + C_p} \frac{g_m}{C_{load}} t_{phase}. \tag{4.44}$$

The effective input-referred mean square noise voltage is obtained by dividing the mean square noise at the output by the square of the system gain. The root mean square (RMS) value of the input-referred noise is

$$v_{i,n,rms} = \frac{2C + C_p}{2C} \sqrt{\frac{S_0}{t_{phase}}}. \tag{4.45}$$

This noise is only 6.9 pV, and can therefore be ignored.

The noise model for the first sampling operation is shown in Fig. 4-18. The total noise is the sum of thermal noise in the op-amp and switch noise when the sampling switches open. The op-amp thermal noise referred to the ADC input is

$$v_{n,1,oa}^2 = S_0 \cdot \mathrm{BW(Hz)} \frac{\pi}{2} \left(\frac{2C + C_p}{2C} \right)^2 = \frac{1}{4} S_0 \frac{g_{m2}}{C_{out} + 2C + C_p} \left(\frac{2C + C_p}{2C} \right)^2, \tag{4.46}$$

which has an approximate rms value of 115 μV.

The switch noise is the portion of the thermal noise from the switches which appears across the sampling capacitors. The noise spectrum produced by the switches is shown in Fig. 4-17. The op-amp suppresses the portion of the noise below ω_1, the op-amp unity-gain bandwidth. The cutoff frequency for the total switch thermal noise is ω_2. These frequencies are

$$\omega_1 = \frac{g_{m,i}}{C_{load}} \tag{4.47}$$

$$= \frac{g_{m,i}}{C_{out} + 2C + C_p} \tag{4.48}$$

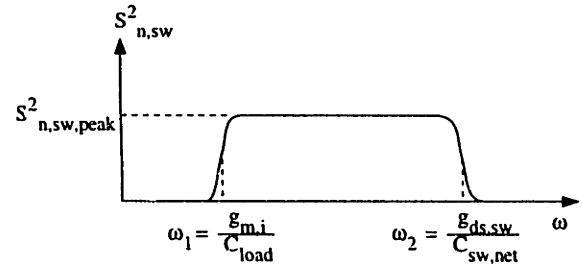


Figure 4-17: Input-referred noise spectrum for switch resistance.

$$= 110\text{Mrad/s} \tag{4.49}$$

$$\omega_2 = \frac{g_{ds,sw}}{C_{net,sw}} \tag{4.50}$$

$$= \frac{g_{ds,sw}}{(2C + C_p)||C_{out}} \tag{4.51}$$

$$= 1200\text{Mrad/s}, \tag{4.52}$$

so the switch bandwidth is significantly greater than the op-amp bandwidth. The effect of the op-amp is ignored in the following analyses, leading to a small overestimate of switch noise.

The total mean square switch noise is

$$v_{n,sw,tot}^2 = 2 \cdot \frac{kT}{C_{out}||(2C + C_p)}, \tag{4.53}$$

where the factor of two accounts for the two switches, and the capacitance used is the net capacitance for each switch. The fraction of the total switch noise which appears at the inverting op-amp terminal is

$$v_{n,sw,in-}^2 = v_{n,sw,tot}^2 \left(\frac{C_{out}}{C_{out} + 2C + C_p} \right)^2. \tag{4.54}$$

The noise at the op-amp inverting input can be referred back to the input by taking into account the capacitive divider between v_{in} and the op-amp inverting input.

$$v_{n,sw,in}^2 = v_{n,sw,in-}^2 \left(\frac{2C + C_p}{2C} \right)^2. \tag{4.55}$$

Combining these equations gives

$$v_{n,sw,in}^2 = 2 \cdot \frac{kT}{C_{out}||(2C + C_p)} \left(\frac{C_{out}}{C_{out} + 2C + C_p} \right)^2 \left(\frac{2C + C_p}{2C} \right)^2 \tag{4.56}$$

$$= 2 \cdot kT \frac{C_{out} + 2C + C_p}{C_{out}(2C + C_p)} \left(\frac{C_{out}}{C_{out} + 2C + C_p} \right)^2 \left(\frac{2C + C_p}{2C} \right)^2 \tag{4.57}$$

$$= \frac{kT}{C} \frac{2C + C_p}{2C} \frac{C_{out}}{C_{out} + 2C + C_p}. \tag{4.58}$$

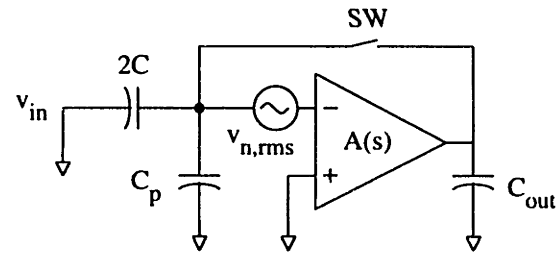


Figure 4-18: Noise sampling operation for first sample (single-ended).

The rms switch noise referred back to the stage input is approximately 74 μV. The total rms noise for the first sampling operation is the sum of the op-amp thermal noise and switch noise, which is 137 μV.

The input sampling noise for subsequent samples can be estimated from Fig. 4-19. Three noise sources contribute to the sampling noise: thermal noise in the first op-amp, thermal noise in the second op-amp, and switch noise when the sampling switches open. Noise on the sampling capacitor 2C due to thermal noise in the first op-amp is

$$\begin{aligned} v_{n,1}^2 &= (S_0 \mathrm{UGBW(Hz)} \frac{\pi}{2}) \left(\frac{2C + C_p}{C} \right)^2 \\ &= \left(\frac{1}{4} S_0 \frac{g_m}{C_{out} + 2C + C||(C + C_p)} \frac{C}{2C + C_p} \right) \left(\frac{2C + C_p}{C} \right)^2 \\ &= \frac{1}{4} S_0 \frac{g_m}{C_{out} + 2C + C||(C + C_p)} \frac{2C + C_p}{C}, \end{aligned}$$

where UGBW is the op-amp unity-gain bandwidth and S_0 is the input-referred thermal noise spectral density (V^2/Hz). The rms noise is approximately 143 μV^2. All of the noise is assumed to appear across the sampling capacitors because the second op-amp maintains a virtual ground at its input.

Sampling noise due to thermal noise in the second op-amp is

$$\begin{aligned} v_{n,2,lowf}^2 &= \left(S_0 \mathrm{UGBW(Hz)} \frac{\pi}{2} \right) \left(\frac{2C + C_p}{2C} \right)^2 \\ &= \frac{1}{4} S_0 \frac{g_{m2}}{C_{out} + 2C + C_p} \left(\frac{2C + C_p}{2C} \right)^2, \end{aligned}$$

which has an estimated rms value of 115 μV. This calculation assumes that the output of the first op-amp is a virtual ground, which is only true at low frequency. As frequency increases, the gain of the first op-amp falls while the gain of the second op-amp is still high, due to the

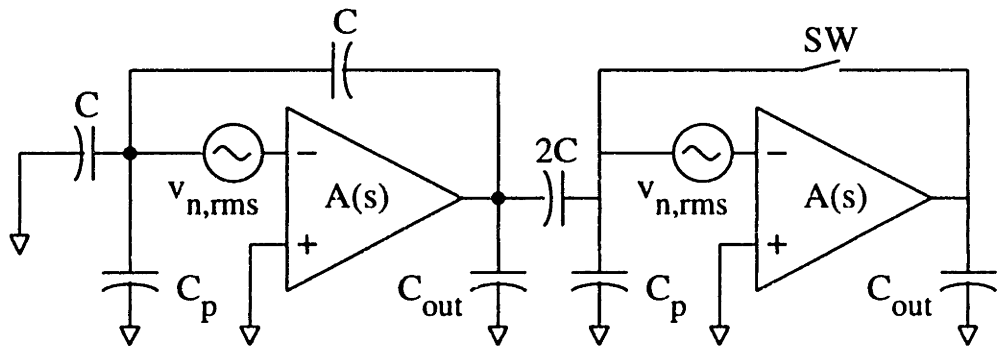


Figure 4-19: Noise sampling operation for subsequent samples (single-ended).

different closed-loop gains. In the extreme case where the gain of the first op-amp is zero, the effective output impedance of the first stage is just the output capacitance.

$$v_{n,2,highf}^2 = \frac{1}{4} S_0 \frac{g_{m2}}{C_{out} + C_p + 2C||(C_{out} + C||(C + C_p))} \left(\frac{C_{out} + C||(C + C_p)}{2C + C_{out} + C||(C + C_p)} + \frac{C_p}{2C} \right)^2 \tag{4.59}$$

The estimated rms noise for the high frequency case is 83.8 μV. The actual noise is between the low frequency and high frequency estimates.

Switch SW produces kT/C_{net} noise. The series combination of the capacitances on the right and left of the switch is

$$C_{net} = C_{out}||(C_p + 2C||(C_{out} + C||(C + C_p))). \tag{4.60}$$

The fraction of the switch noise which appears across the sampling capacitor can be calculated by assuming that the first op-amp does not respond while the switch opens, which is reasonable considering the relative speeds of the op-amp and the switch. The reduction of switch noise at low frequency due to op-amp feedback is ignored for the reason stated previously, namely, that the switch bandwidth is much greater than the op-amp bandwidth. The net switch noise is

$$v_{n,sw}^2 = \frac{2kT}{C_{net}} \left(\frac{C_{out}}{C_{out} + C_p + 2C||(C_{out} + C||(C + C_p))} \right)^2 \left(\frac{C_{out} + C||(C + C_p)}{2C + C_{out} + C||(C + C_p)} + \frac{C_p}{2C} \right)^2, \tag{4.61}$$

which is approximately 71 μV.

The total noise is found by summing the thermal noise contributions of the op-amps and the switch noise.

$$v_{n,rms} = \sqrt{v_{n,1}^2 + v_{n,2}^2 + v_{n,sw}^2} \tag{4.62}$$

The total rms noise for each sampling operation except the first is 197 μV.

In terms of the residue plot, random noise adds a band of uniform width around the ideal residue plot. Noise can be input-referred by adding the noises introduced at each sampling operation.

$$v_{n,rms}^2 = (137\mu\text{V})^2 + (197\mu\text{V})^2 \left(\frac{1}{4} + \frac{1}{16} + \frac{1}{64} + \cdots\right) \quad (4.63)$$

$$v_{n,rms} = 195\mu\text{V} \quad (4.64)$$

Although the total rms noise was calculated assuming an infinite number of bits, the noise is not significantly different for an 8b or 10b conversion. The noise is well below the 2 mV lsb level.

4.1.4 CDS Circuit

Figure 4-20 shows a complete schematic diagram of the CDS circuit; the associated timing diagram is Fig. 4-22. The CDS circuit reduces fixed-pattern noise by subtracting the pixel output at reset from the pixel output at the end of the integration period. It also converts the single-ended output from the pixel to a fully differential voltage.

When a pixel reaches the end of its integration period, its row select turns on. At the same time, **sample1** goes high. The CDS circuit configuration during this interval is shown in Fig. 4-21(a). After the column line voltage $v_{line}(t_1)$ and op-amp output V_o have settled to the required accuracy, **sample1** goes low. Feedback switches M9 and M10 are timed to open slightly before the other switches clocked on **sample1** to reduce clock feedthrough. Next, **sample2** goes high and **clkb** goes low. There is no specific delay required between the falling edge of **sample1** and the rising edge of **sample2**, but the two clocks must not overlap. The falling edge of **clkb** advances the barrier generation shift register so that the barrier gate voltage rises to the reset level. The CDS configuration during this period is shown in Fig. 4-21(b). After the column line voltage $v_{line}(t_2)$ and the op-amp output have settled, **isolate** goes low, isolating the CDS circuit from any change in potential on the column line. This configuration is shown in Fig. 4-21(c).

The resultant output voltage is found by equating the charge stored on the summing nodes before and after the **sample1** phase. The output voltage at the end of the **sample2** phase is

$$V_o = (v_{cm} - v_{os}) + (v_{line}(t_2) - v_{line}(t_1)). \quad (4.65)$$

The common-mode voltage v_{cm} is 2 V, the offset voltage v_{os} is 3 V, and the column line reset voltage $v_{line}(t_2)$ is approximately 2.5 V. The pixel output voltage at the end of the integra-

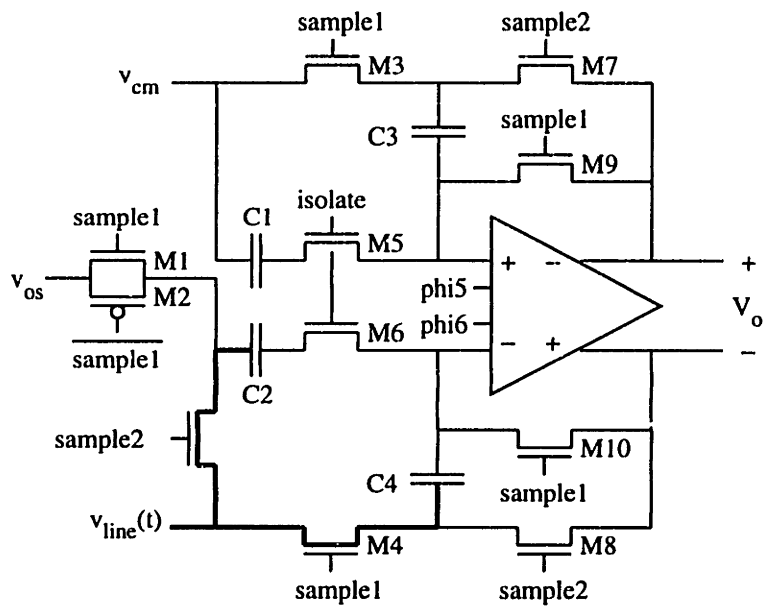


Figure 4-20: Complete CDS circuit schematic. Input signal path is shown in bold. On sample phase, signal is passed onto C4. On reset phase, signal is passed onto C2.

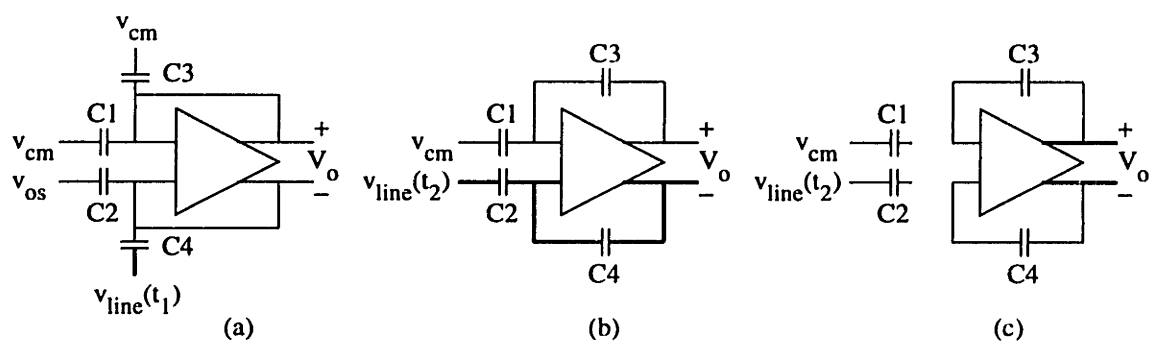


Figure 4-21: CDS clocking sequence. The signal path is shown in bold.

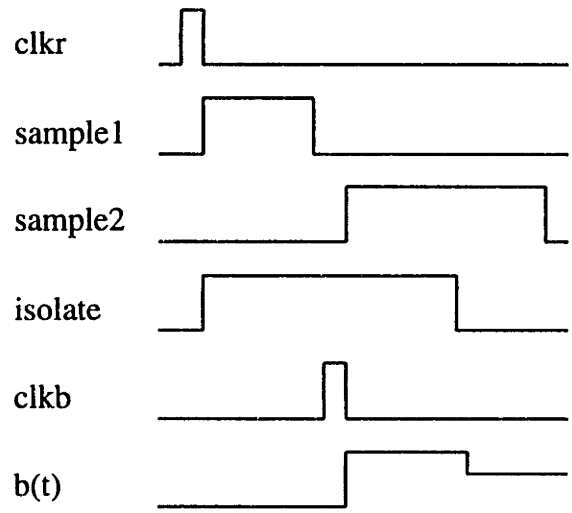


Figure 4-22: CDS circuit timing diagram.

tion period varies from about 2.5 V to 0.5 V, so that the total output voltage ranges from approximately -1 V to +1 V.

Errors in CDS Circuit

The CDS operation is nonideal due to random noise, capacitor mismatch, op-amp input offset voltage, finite op-amp gain, op-amp CMRR, and clock feedthrough.

Random Noise Since correlated double-sampling removes $1/f$ noise to first order, only thermal noise will be examined. Noise arises from three operations. When the **sample1** switches open, switch noise and op-amp input-referred noise are sampled at the summing node. The noise model for this operation is shown in Fig. 4-23(a). When the **isolate** switches open, switch and op-amp noise are again sampled at the summing node, as shown in Fig. 4-23(b). When the ADC samples the CDS output, direct noise is present due to the op-amp noise, as shown in Fig. 4-23(c).

The sampled noise in Fig. 4-23(a) is the sum of the switch noise $v_{n,sw1}$ and the op-amp input-referred noise $v_{n,oa,1}$. The effect of the op-amp on the switch noise is ignored, as usual, because the switch bandwidth is much larger than the op-amp bandwidth. The mean square noise charge trapped on the summing node is

$$q_{n,1}^2 = \frac{2kT}{C_{out}||(2C+C_p)}\left(\frac{C_{out}}{C_{out}+2C+C_p}\right)^2(2C+C_p)^2 + S_0 \cdot \text{UGBW(Hz)} \cdot \frac{\pi}{2}(2C+C_p)^2. \quad (4.66)$$

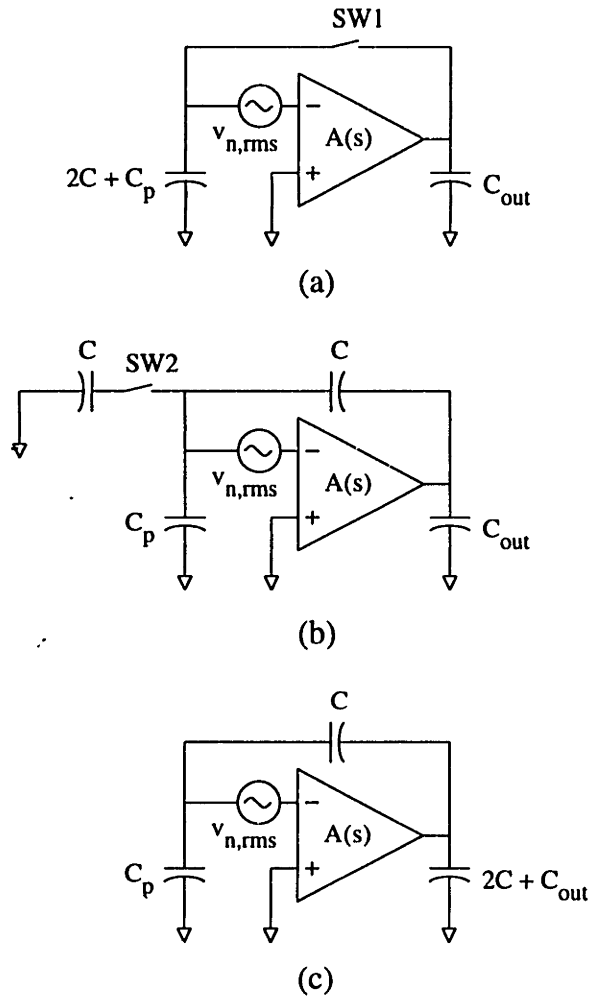


Figure 4-23: (a) CDS noise model when **sample1** switches open. (b) CDS noise model when **isolate** switches open. (c) CDS noise model for direct noise.

Referred to the CDS input, the mean square noise noise voltage is

$$v_{n,in,1}^2 = \frac{2kT}{C_{out}||(2C+C_p)}\left(\frac{C_{out}}{C_{out}+2C+C_p}\right)^2\left(\frac{2C+C_p}{C}\right)^2 + \frac{1}{4}S_0\frac{g_{m2}}{C_{out}+2C+C_p}\left(\frac{2C+C_p}{C}\right)^2 \tag{4.67}$$

When the **isolate** switches open, the CDS noise model is as shown in Fig. 4-23(b). The noise charge trapped on the summing node is identical in magnitude to the noise charge on the input capacitor. Referred to the CDS input, the mean square noise voltage is therefore the noise voltage on the input capacitor.

$$v_{n,in,2}^2 = \frac{kT}{C||(C_p+C||C_{out})}\left(\frac{C_p+C||C_{out}}{C+C_p+C||C_{out}}\right)^2 + \frac{1}{4}S_0\frac{g_{m2}}{C_{out}+C||(C+C_p)}\left(\frac{C}{2C+C_p}\right). \tag{4.68}$$

The direct noise is the op-amp noise sampled by the ADC. The noise model for direct noise is shown in Fig. 4-23(c). The input-referred and output-referred noises are identical because the CDS has unity gain, so the mean square input noise is

$$v_{n,in,3}^2 = \frac{1}{4}S_0\frac{g_{m2}}{2C+C_{out}+C||C_p}\left(\frac{C+C_p}{C}\right). \tag{4.69}$$

The total CDS input-referred noise is

$$v_{n,in,rms} = \sqrt{v_{n,in,1}^2 + v_{n,in,2}^2 + v_{n,in,3}^2}, \tag{4.70}$$

which is approximately 308 μV. This assumes that the values of C, C_p, and C_{out} are the same for the CDS circuit as for the ADC, which is roughly valid.

Capacitor Mismatch Capacitors C1 - C4 have the same nominal value C. Mismatches in these capacitances produce gain and offset errors. These errors can be derived by calculating the output voltage v_o as a function of those capacitances and taking partial derivatives with respect to each capacitance.

The output voltage is determined by the condition that charge on the summing nodes is the same in the **sample1** and **sample2** phases.

$$\begin{aligned} q_{in+} &= (v'-v_{cm})(C1+C3)+v'C_{p+} && (4.71)\\ &= (v''-v_{o-})C3+(v''-v_{cm})C1+v''C_{p+} && (4.72)\\ q_{in-} &= (v'-v_{os})C2+(v'-v_{line}(t_1))C4+v'C_{p-} && (4.73)\\ &= (v''-v_{o+})C4+(v''-v_{line}(t_2))C2+v''C_{p-}, && (4.74) \end{aligned}$$

where v_{os} is the offset voltage, $v_o = v_{o+} - v_{o-}$ is the final output voltage, $v' = v_{cm}$ is the voltage at each op-amp input during the **sample1** phase, C_{p+} and C_{p-} are the parasitic capacitances at the positive and negative op-amp input terminals, and v'' is the voltage at each op-amp input during the **sample2** phase. The simultaneous solution of these equations is

$$v_o = \frac{2(C1 + C3 + C_{p+})}{C3(C2 + C4 + C_{p-}) + C4(C1 + C3 + C_{p+})}(v_{line}(t_1)\cdot C4 - v_{line}(t_2)\cdot C2 + v_{os}\cdot C2 - v_{cm}\cdot C4). \tag{4.75}$$

Sensitivities to capacitor mismatch are found by taking partial derivatives of v_o with respect to each capacitance. The sensitivities are

$$\frac{\Delta v_o}{\Delta C1/C1} = \frac{C}{2(2C + C_p)} v_{o,nom} \tag{4.76}$$

$$\frac{\Delta v_o}{\Delta C2/C2} = (v_{os} - v_{line}(t_2)) - \frac{C}{2(2C + C_p)} v_{o,nom} \tag{4.77}$$

$$\frac{\Delta v_o}{\Delta C3/C3} = -\frac{C + C_p}{2(2C + C_p)} v_{o,nom} \tag{4.78}$$

$$\frac{\Delta v_o}{\Delta C4/C4} = (v_{line}(t_2) - v_{os}) + \frac{C + C_p}{2(2C + C_p)} v_{o,nom} \tag{4.79}$$

$$\frac{\Delta v_o}{\Delta C_{p+}/C_{p+}} = \frac{C_p}{2(2C + C_p)} v_{o,nom} \tag{4.80}$$

$$\frac{\Delta v_o}{\Delta C_{p-}/C_{p-}} = -\frac{C_p}{2(2C + C_p)} v_{o,nom}, \tag{4.81}$$

where $v_{o,nom}$ is the nominal output voltage. The resultant offset and gain errors are

$$\text{offset error} = \left(\frac{\Delta C2}{C2} - \frac{\Delta C4}{C4}\right)(v_{line}(t_2) - v_{os}) \tag{4.82}$$

$$\text{gain error} = \frac{v_{o,nom}}{2(2 + C_p/C)}\left(\frac{\Delta C1}{C1} - \frac{\Delta C2}{C2} + (1 + C_p/C)\left(\frac{\Delta C4}{C4} - \frac{\Delta C3}{C3}\right) + \frac{C_p}{C}\left(\frac{\Delta C_{p+}}{C_{p+}} - \frac{\Delta C_{p-}}{C_{p-}}\right)\right). \tag{4.83}$$

Using the usual approximations for C_p and C, and assuming that these capacitors match to 0.2%, the rms offset error is 1.4 mV (0.71 lsb) and the rms gain error is 0.053%. Although channel-to-channel accuracy is limited by capacitor matching in the CDS and ADC, dynamic range is unaffected.

Input Offset Voltage and Finite Op-Amp Gain Input-referred offset voltage is completely cancelled when the op-amp gain is infinite. If the gain is finite, the actual output

voltage can be found as before by equating the charges stored at the op-amp inputs. In the **sample1** phase,

$$q_{in+} = 2(v_{o-} - v_{cm})C + v_{o-}C_p \quad (4.84)$$

$$= \frac{A}{1+A}v_{os}C + \left(v_{cm} + \frac{1}{2}\frac{A}{1+A}v_{os}\right)C_p \quad (4.85)$$

$$q_{in-} = (2v_{o-} - v_{offset} - v_{line}(t_1))C + v_{o-}C_p \quad (4.86)$$

$$= \left(2v_{cm} - \frac{A}{1+A}v_{os} - v_{offset} - v_{line}(t_1)\right)C + \left(v_{cm} - \frac{1}{2}\frac{A}{1+A}v_{os}\right)C_p. \quad (4.87)$$

In the **sample2** phase,

$$q_{in+} = (2v_{in+} - v_{o-} - v_{cm})C + v_{in+}C_p \quad (4.88)$$

$$q_{in-} = (2v_{in-} - v_{o+} - v_{line}(t_2))C + v_{in-}C_p. \quad (4.89)$$

By the definition of offset voltage,

$$v_o = v_{o+} - v_{o-} = A(v_{in+} - v_{in-} - v_{os}). \quad (4.90)$$

Equating $q_{in+} - q_{in-}$ in the **sample1** and **sample2** phases gives

$$v_o = \frac{\left(-\frac{2}{1+A}v_{os} + v_{offset} - v_{cm} + v_{line}(t_1) - v_{line}(t_2)\right)C - \frac{1}{1+A}v_{os}C_p}{\left(1+\frac{2}{A}\right)C + \frac{1}{A}C_p}, \quad (4.91)$$

which can be approximated by

$$v_o \approx v_{o,nom} - \frac{1}{A}(v_{o,nom} + v_{os})\left(2 + \frac{C_p}{C}\right). \quad (4.92)$$

Finite op-amp gain and op-amp input offset voltage produce both a gain and an offset error. Because the op-amp gain is high (20,000), gain error due to finite op-amp gain is only about 0.01% and offset error is unlikely to exceed 5 μV.

CMRR One problem with the design of the CDS circuit is the shift in input common mode between the two configurations. This is unavoidable because the pixel output voltages during the **sample1** and **sample2** phases are not symmetric about v_{cm}. Two undesireable effects result. Mismatch in the parasitic capacitances C_p at the two inputs causes differential charge injection into the input nodes when the input common mode changes. The resulting change in output voltage has already been derived. The other effect is that device mismatches in the op-amp produce a differential output in response to a change in the input common mode voltage.

This can be modeled by an effective differential input signal $\Delta v_{i,dm}$, where

$$\Delta v_{i,dm} = \frac{\Delta v_{i,cm}}{\text{CMRR}}, \text{ where} \tag{4.93}$$

$$\Delta v_{i,cm} = \frac{C}{2(2C + C_p)} v_o. \tag{4.94}$$

The effective change in input differential signal is multiplied by the closed-loop gain of the system, producing a gain error

$$\Delta v_o = \frac{1}{2} \frac{v_o}{\text{CMRR}}. \tag{4.95}$$

Since the CMRR is estimated at 65 dB, the gain error due to finite CMRR is approximately 0.03%.

Clock Feedthrough Clock feedthrough occurs when switching between the **sample1** and **sample2** phases. The magnitude of feedthrough will be the same as for the ADC, since the circuit configuration is identical.

4.1.5 Output Multiplexer and Pad Driver

The output multiplexer schematic is shown in Fig. 4-24. The multiplexer takes eight outputs from four converters and selects one at a time to get buffered and sent to the output pin. The multiplexer latches the ADC outputs, since the outputs are valid only for the duration of an individual clock phase. The corresponding timing diagram is shown as Fig. 4-25. The signals **swsel1**, **swsel2**, and **swsel3** come in from off-chip and generate **switch1** - **switch8** on-chip. The pad driver is a four-stage tapered buffer shown in Fig. 4-26. It drives a 20 pF load with a simulated propagation delay $t_{PLH} = 2$ ns.

4.2 Row Select and Barrier Generation

The barrier generation circuit applies a stepped voltage waveform to the lateral overflow gates in each row. Every row sees the same waveform, but the waveforms are staggered in time between rows. This is neccessary because the pixel does not contain a storage region, and therefore must be read out immediately at the end of its integration period. Since finite time is required to read out the row, the next row's integration period must end slightly later.

One way to produce a staggered barrier waveform is an analog delay line. Errors in the transmission of analog voltage levels are cascaded, causing the barrier levels at the end of the

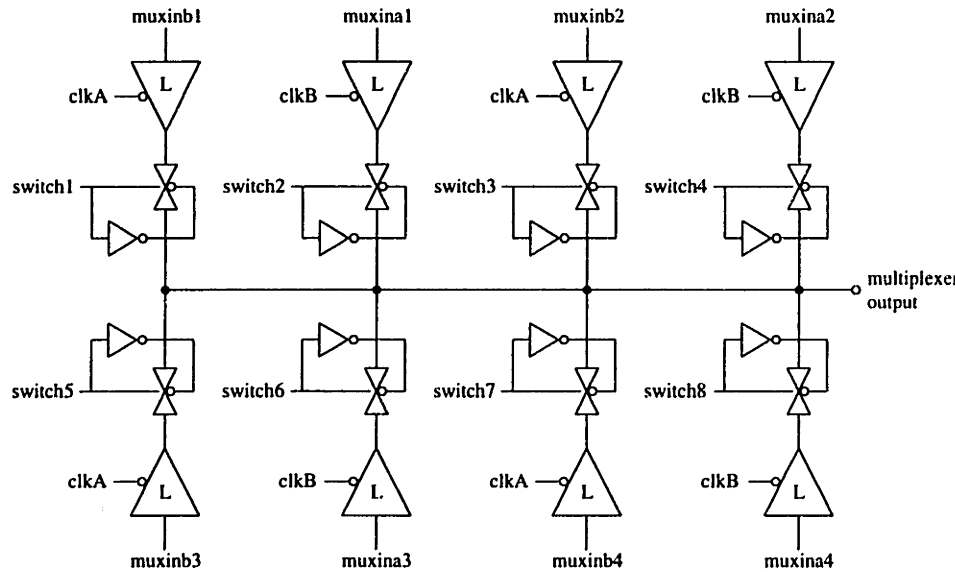


Figure 4-24: Output multiplexer schematic.

delay line to be significantly different from the barrier levels at the start. The imager would be likely to have large row pattern noise.

The method used on this chip is a digital shift register. Bits can be easily transmitted without error down the length of the array. Eight bits are used to connect the lateral overflow gate in each row to one of eight analog voltage levels. The analog voltages are supplied from off-chip, although they could also be generated on-chip. For example, a string of diode-connected PMOS transistors, connected between analog power and ground, will evenly divide the 5 V supply voltage. Step voltages generated in this way will have poor power supply rejection. The idea could be improved by using a bandgap reference to generate the upper voltage. The levels are completely adjustable, and the time points at which the barrier level can change are defined by the falling edge of the shift register clock, which occurs approximately every 4 μs when operated at 1000 fps. Advantages of digitally-controlled compression include good matching between compression and expansion curves and easy and accurate modification of the compression characteristic [53].

Figure 4-27 shows two bits of the barrier generation shift register for one row. The analog voltages are V_1 and V_2. The two-phase non-overlapping clock is **p1** and **p2**, which is generated on-chip from the external clock **clkb**. If d_i is 1, row(i) is connected to V_i when **p2** goes high. If d_i is zero, no connection is made between row(i) and V_i. In each row, exactly one of the digital bits d_i will be a one, so the row is connected to exactly one of the analog levels V_i. The

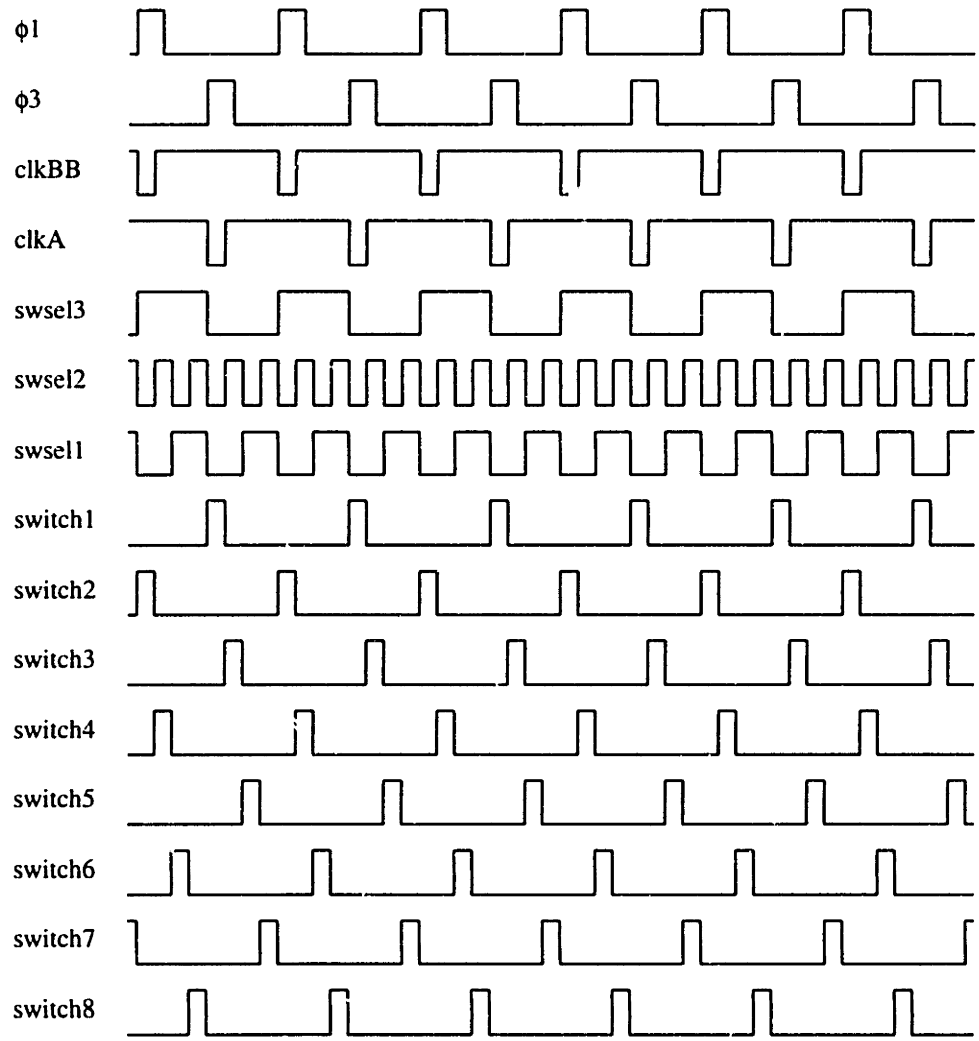


Figure 4-25: Timing diagram for output multiplexer.

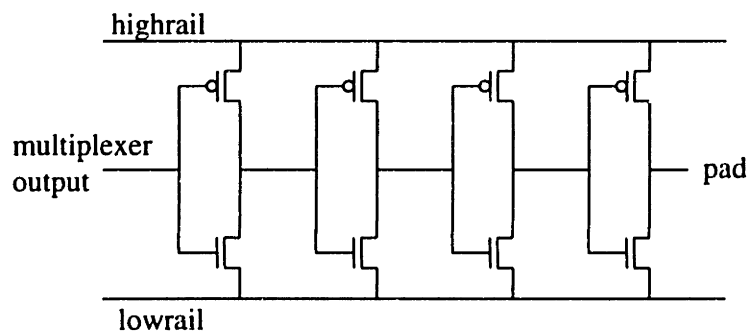


Figure 4-26: Pad driver schematic.

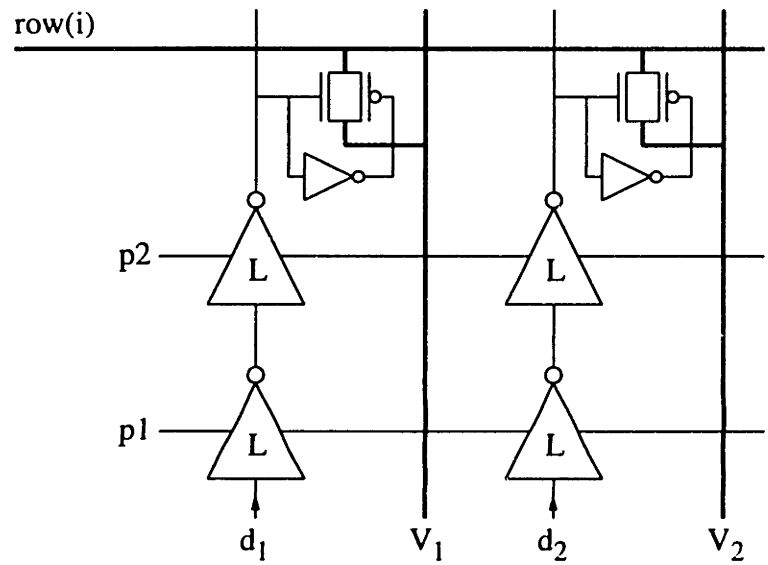


Figure 4-27: Two bits of the the barrier generation circuit for one row.

bit travels to the output unchanged (since the two inversions cancel), where it is passed on to the next row.

Figure 4-28 shows the barrier waveforms for each row and the switch connections at the instant of time indicated by the dashed line. The figure is drawn for an 8-row, 4-step barrier function imager, but this example scales to the actual 256-row, 8-step barrier function test chip. Each time the barrier shift register advances, the dashed line moves to the left by the row time T_r. After eight row times, all rows have been read out and the cycle repeats, so for this example the frame time $T_f = 8\ T_r$. The actual test chip has 256 rows, so $T_f = 256\ T_r$.

The clock generator takes the input clocks for the barrier and row select shift registers, and creates the two-phase nonoverlapping clocks required for their operation. The schematic for each of the clock generators is shown in Fig. 4-29 [54]. Assume that the input clock ϕ is high. The output of the top NOR gate, A, must be low. After propagating through the top inverter chain, ϕ_1 goes low. The bottom NOR gate has two low inputs, and therefore its output B goes high. After propagating through the bottom inverter chain, ϕ_2 goes high. The outputs remain in this state until ϕ goes low. Then B goes low, and after propagating through the bottom inverter chain, ϕ_2 goes low. Now both inputs to the top NOR gate are low, and A goes high. After propagating through the upper inverter chain, ϕ_1 also goes high. Note that only one output can be high at any given time, and when that output goes low, the other output cannot go high until the signal has propagated through a NOR gate and the inverter chain.

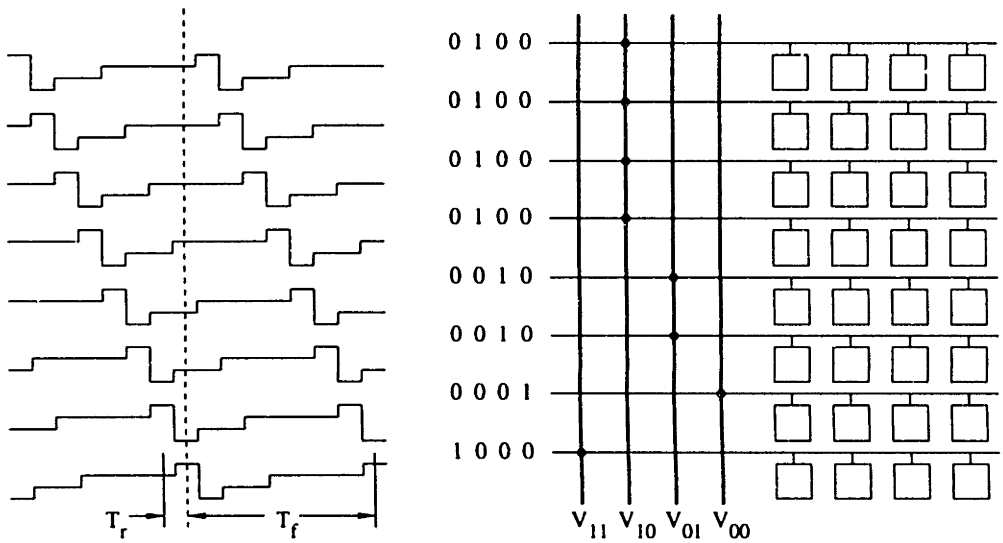


Figure 4-28: Staggered readout example for an 8-row, 4-step barrier function imager. T_r is the row time and T_f is the frame time. The digital word next to each row specifies which analog level V_{ij} is connected to the barrier in that row.

The nonoverlapping period is the sum of the NOR gate delay and six inverter gate delays.

4.3 Pixel Array

The pixel array is a 256 × 256 grid of wide dynamic range pixels. The power supply V_{dd} and the spill gate voltage **cgbias** are common to all pixels in the array. The lateral overflow gate voltage and row select are common to all pixels in a given row, and the the pixel outputs are common to all pixels in a given column. Each column is biased with a 20 μA cascoded current

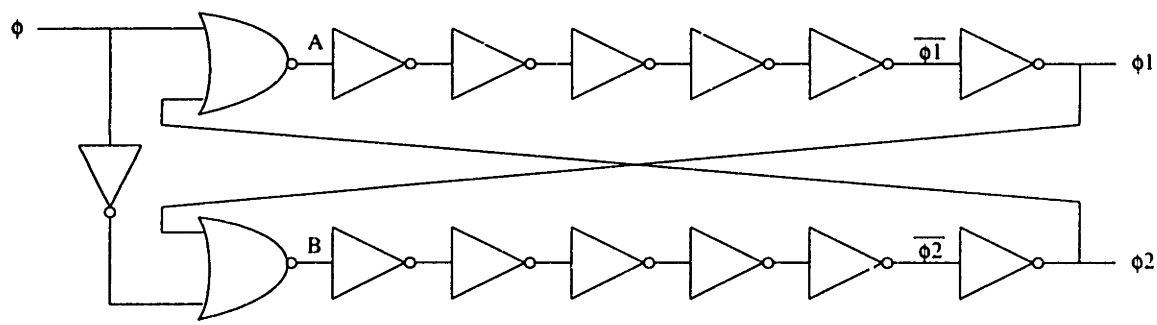


Figure 4-29: Nonoverlapping clock generator.

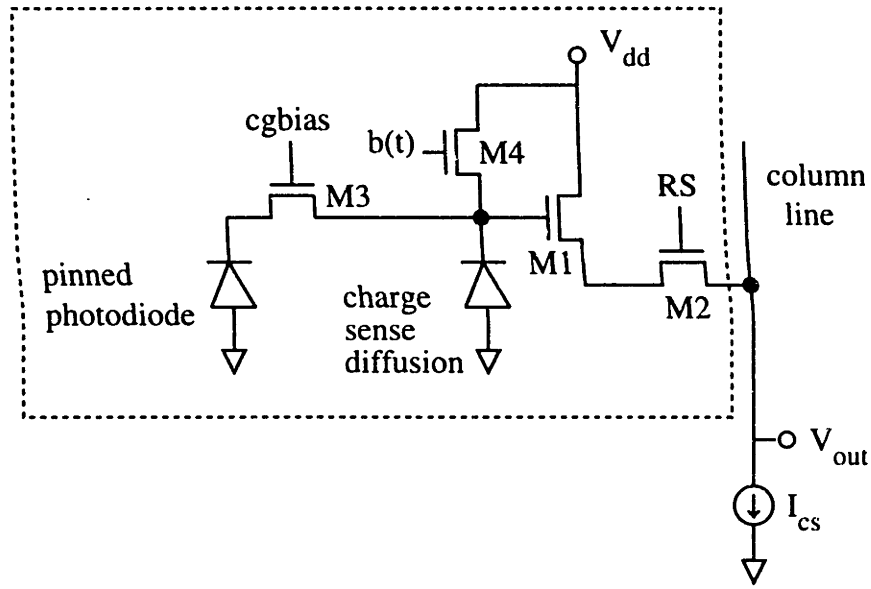


Figure 4-30: Pixel equivalent circuit.

sink to drive the in-pixel source-follower.

4.3.1 Wide Dynamic Range Pixel

The pixel design was described in a previous chapter. The pixel equivalent cicuit is repeated below as Fig. 4-30. In this section, the expected performance characteristics for the test chip are calculated.

Sensitivity

The pixel sensitivity is defined in this context as the ratio of output voltage to input light power.

$$S = \frac{RT_{int}G_{sf}}{C_{sense}}, \tag{4.96}$$

where R is the responsivity (A/W), G_{sf} is the source follower gain, and C_{sense} is the total capacitance at the charge sense node. It is measured by illuminating the array with light of a known power density (W/m^2), measuring the average output code, and subtracting the average output code in the dark. The code can be referred back to pixel output voltage since the reference voltage levels used in the ADC are known.

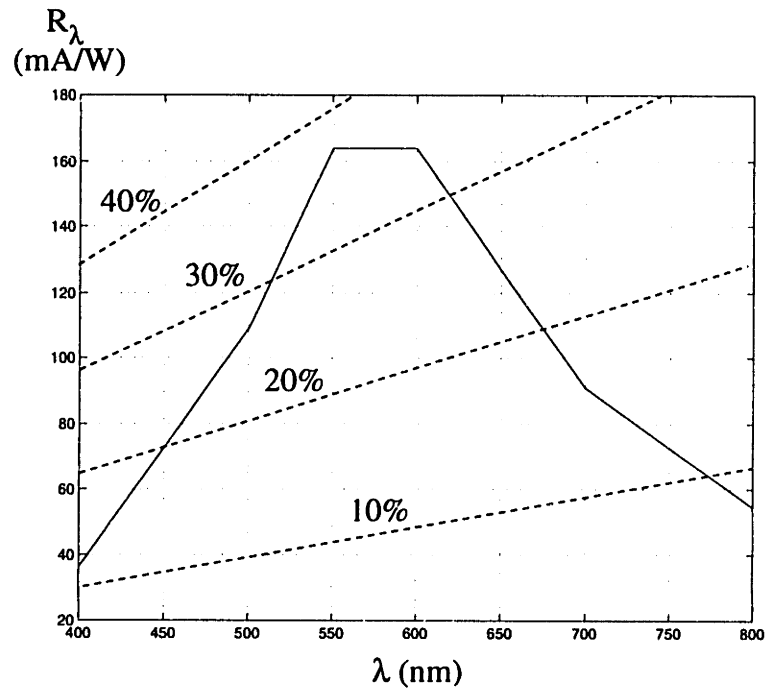


Figure 4-31: Spectral responsivity for n^+-p photodiode. Dashed lines indicate lines of constant quantum efficiency.

Responsivity Responsivity depends on the fraction of the pixel area occupied by diffusion, the fraction of pixel area occupied by windows, and details of the process such as diffusion depth, doping profiles, and overglass thicknesses. Responsivity also has a strong wavelength dependence. Estimated spectral responsivity for an n^+ photodiode in p-substrate is plotted in Fig. 4-31 [55].

Responsivity for a non-monochromatic light source is a weighted average of the responsivity over the relevant bandwidth. For example, a blackbody radiation source at absolute temperature T has a spectrum

$$B(\lambda) = \frac{2\pi hc^2}{\lambda^5} \frac{1}{\exp(hc/\lambda kT) - 1}. \tag{4.97}$$

The net responsivity over the visible spectrum is

$$R = \frac{\int_{400nm}^{800nm} B(\lambda) R_\lambda d\lambda}{\int_{400nm}^{800nm} B(\lambda) d\lambda}. \tag{4.98}$$

Based on the spectral responsivity in Fig. 4-31, the responsivity to blackbody radiation at T = 2856 K is approximately 103 mA/W.

Source Follower Gain The source follower gain is

$$G_{sf} = \frac{g_{m1}}{g_{m1} + g_{mb1} + g_{o1} + g_{cs}}, \tag{4.99}$$

where g_{m1}, g_{mb1}, and g_{o1} are the transconductance, back-gate transconductance, and drain-source conductance, respectively, of the source follower device M1 in Fig. 4-30. The output conductance of the current source is g_{cs}; this term is usually negligible compared to the other terms. The simulated source follower gain is a weak function of output voltage, ranging from 0.792 to 0.889 over the pixel output range.

Sense Capacitance The sense capacitance C_{sense} is found by summing all of the capacitances at the charge sense node.

$$\begin{aligned} C_{sense} = {} & A_{M1-sub}C_{A,M1-sub} + P_{M1-sub}C_{P,M1-sub} + A_{M1-poly}C_{A,M1-poly} + \\ & P_{M1-poly}C_{P,M1-poly} + A_{poly-sub}C_{A,poly-sub} + P_{poly-sub}C_{P,poly-sub} + \\ & A_{diff}C_{A,diff}(0V) + P_{diff}C_{P,diff}(0V) + 3W_{ol}C_{ol} + A_{sf}C_{gate}(1 - G_{sf}), \end{aligned} \tag{4.100}$$

where

- A_{M1-sub} is the area of the metal1 - substrate parasitic capacitance
- $C_{A,M1-sub}$ is the parasitic area capacitance between metal1 and substrate
- P_{M1-sub} is the perimeter of the metal1 - substrate parasitic capacitance
- $C_{P,M1-sub}$ is the parasitic perimeter capacitance between metal1 and substrate
- $A_{M1-poly}$ is the area of the metal1 - poly parasitic capacitance
- $C_{A,M1-poly}$ is the parasitic area capacitance between metal1 and poly
- $P_{M1-poly}$ is the perimeter of the metal1 - poly parasitic capacitance
- $C_{P,M1-poly}$ is the parasitic perimeter capacitance between metal1 and poly
- $A_{poly-sub}$ is the area of the poly - substrate parasitic capacitance
- $C_{A,poly-sub}$ is the parasitic area capacitance between poly and substrate
- $P_{poly-sub}$ is the perimeter of the poly - substrate parasitic capacitance

- $C_{P,poly-sub}$ is the parasitic perimeter capacitance between poly and substrate
- A_{diff} is the area of the diffusion - substrate parasitic capacitance
- $C_{A,diff}(0\text{ V})$ is the parasitic area capacitance between diffusion and substrate
- P_{diff} is the perimeter of the diffusion - substrate parasitic capacitance
- $C_{P,diff}(0\text{ V})$ is the parasitic perimeter capacitance between diffusion and substrate with 0 V applied across the junction
- W_{ol} is the width of the transistors abutting the sense diffusion
- C_{ol} is the overlap capacitance per unit length of the abutting transistors
- A_{sf} is the source follower gate area
- C_{gate} is the capacitance per unit area of a transistor gate
- G_{sf} is the source follower gain.

The sense capacitance varies from 9.32 fF to 11.0 fF.

Saturation Level

Pixel capacity is product of sense capacitance and usable voltage range. Taking the nominal output range of the pixel as 2.5 V to 0.5 V, the sense diffusion voltage ranges from 3.87 V to 1.54 V. Using a nominal 10 fF value for the sense capacitance, the saturation charge is approximately 146,000 e^-. Of course, the corresponding saturation illumination depends on the barrier curve.

4.3.2 Settling Time

The column line settling time is determined by the time constant of the column line capacitance and the output conductance of the source follower. The simulated column line capacitance is 3.83 pF, the simulated output conductance $g_{m1} + g_{mb1}$ of the source follower is 213 μS, and the minimum drain-source conductance g_{ds2} of M2 is 1.51 mS, giving a time constant

$$\tau = \frac{C_{line}}{g_{m1} + g_{mb1} || g_{ds2}} = 21 \text{ ns.} \tag{4.101}$$

Adequate settling requires around ten time constants, or 210 ns. The total sampling time for the column voltage is at least 1 μs, so the column settling time is adequate.

Random Noise

Photonic shot noise and barrier noise are the major components of random noise. Shot noise depends on free charge, which is dependent on illumination and the barrier function. The worst case is when the chip is operated as a linear imager, and the pixel just saturates. The shot noise in this case is 382 e^-. The barrier noise is given by

$$q_{n,barr} = \sqrt{2kTC_{sense}} = 56 \text{ e}^-, \tag{4.102}$$

where the factor of two is due to the double-sampling operation. Barrier noise represents 0.04% of the full-scale signal.

4.4 Imager Operation

The complete row timing diagram for 12 bit output is shown in Fig. 4-32. The row select advances on the falling edge of **clkr**. As the outputs from the next selected row become available on the column lines, the CDS circuits are in the **sample1** phase to sample these outputs. After the CDS circuits have settled, **sample1** goes low and **clkb** is strobed, advancing the barrier waveform. The barrier in the selected row goes to the highest potential, resetting the pixels in that row. The **sample2** clock goes high to sample the pixel output values during reset. After the column line voltages settle, **isolate** goes low to hold the sampled pixel outputs. When ϕ**0** goes high, the ADC accepts input from the CDS circuit specified by **Aselect** and **Bselect**. After digitizing this quantity, the ADC accepts input from the other CDS circuit and digitizes that quantity. Outputs from four ADC's are latched into the output multiplexer by **clkA** and **clkBB**, and are selected for output by **swsel1**, **swsel2**, and **swsel3**.

The frame timing diagram is shown in Fig. 4-33. The row select shift register input **rsin** defines the start of the frame. It is high for only one row time, so that only a single "1" is present in the row select shift register. The column select bits **colsel(i)** are held constant during each row time, but their values depend on the choice of barrier function.

4.5 Bias Circuit

The currents in the pixel source followers and op-amps are set by the on-chip bias circuit, shown in Fig. 4-34. The bias circuit takes three current inputs and generates four bias voltages: **oa1bias** for the CDS op-amps, **oa2bias** for the ADC op-amps, and **csbias** and **ccbias** for the pixel source followers. The devices and currents are large in order to generate low impedance

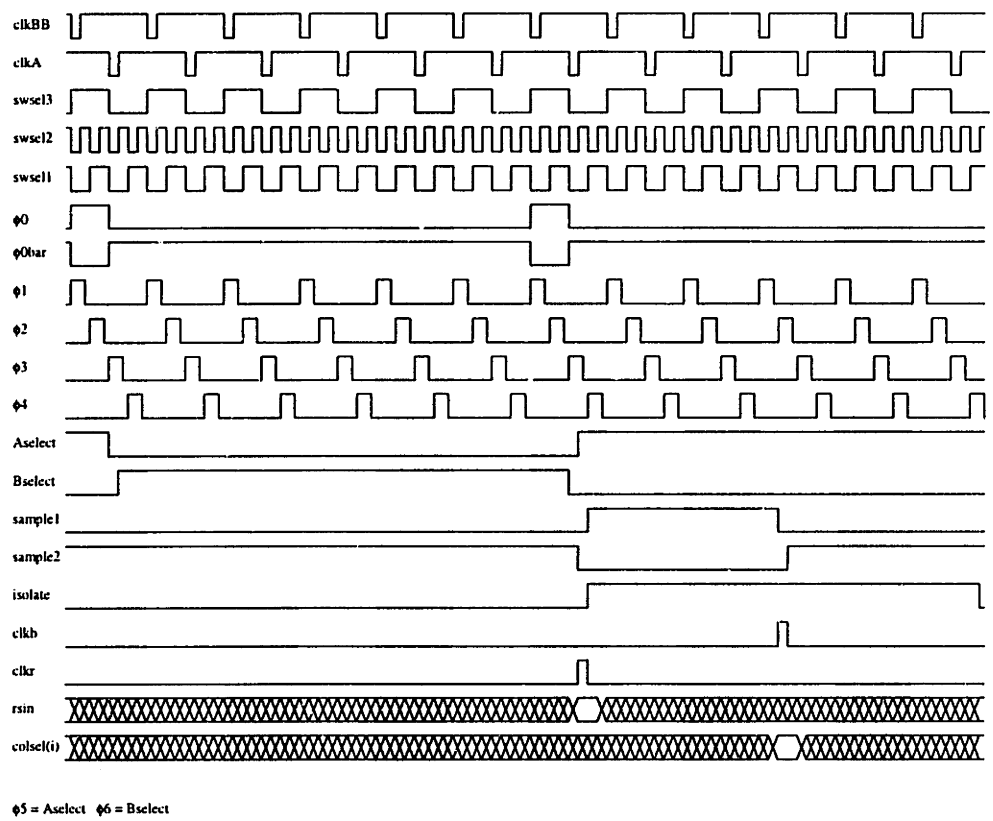


Figure 4-32: Complete timing diagram for one row time.

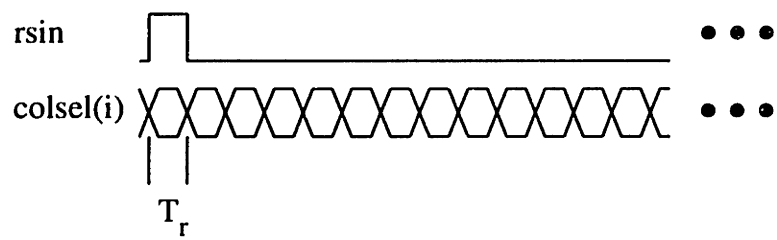


Figure 4-33: Frame timing diagram. Total frame time $T_f = 256\ T_r$.

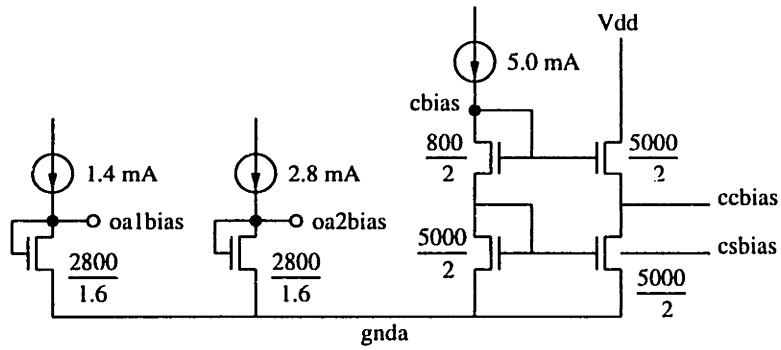


Figure 4-34: Bias circuit for test chip.

bias voltages. This minimizes noise coupling and slow-settling time constants in the transient response caused by changing bias point.

Chapter 5

Experimental Results

The experimental setup consists of a camera board, a test board, the experimental imager chip, a lens and lens mount, light sources, filters, apertures, integrating sphere, and a host PC. The test board can be connected to the camera board, allowing characterization of the on-chip test structures. Frames can be downloaded using only the camera board and imager chip. The camera board has a DB-25 connector which is used to interface to the parallel port on the host PC.

The performance of each major imager component (pixel, CDS, and ADC) was determined from individual test structures. The test circuit, test method, and results are discussed for each of these structures. The transfer characteristic relating input illumination to output digital code is shown for both the linear and compressive imaging modes. Sets of uniformly illuminated frames are used to determine fixed-pattern and random noise. Sample frames contrasting linear and compressive imaging are presented.

5.1 On-Chip Test Structures

5.1.1 Pixel

The equivalent schematic diagram of the on-chip individual test pixel is shown in Fig. 5-1. The test pixel is identical to an array pixel, except that a PMOS source follower is added to buffer the output. Responsivity and conversion gain are measured. The transfer characteristic is determined for a quasi-logarithmic compression function. The quasi-logarithmic compression function is logarithmic over a limited range of illumination; at very low and very high illumination, the response must be linear with illumination. Time-domain output voltage waveforms

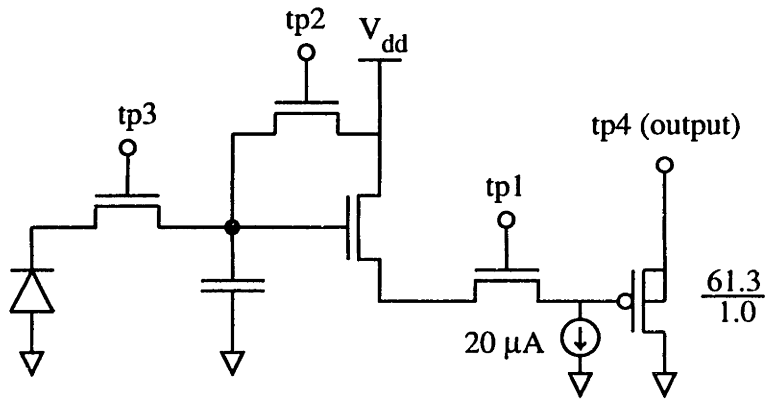


Figure 5-1: Schematic diagram of on-chip individual test pixel.

are shown for low, medium, and high illumination.

The pixel output **tp4** is connected to an off-chip 20 μA current source. The charge spill gate **tp3** is connected to the off-chip 1 V bias source used by the pixel array. The row select line **tp1** is connected to the **rsin** clock, although this signal is usually kept high when performing pixel tests. The barrier gate voltage **tp2** is connected to the output of an off-chip 8-to-1 analog multiplexer. The multiplexer inputs are the bias voltages **col1-8.**

Responsivity

Responsivity cannot be characterized directly from the test pixel since the charge sense node is not directly accessible. This parameter is determined from a 10 × 10 array of photodiodes, each of which is identical to an array pixel except that the photocurrent is taken out directly. Their combined photocurrents are brought off-chip and measured with an HP 4140B picoammeter.

The test setup used to determine responsivity is shown in Fig. 5-2. In Fig. 5-2, the imager chip is placed directly behind a 0.4 mm diameter aperture and current from the photodiode array is measured. Aperture size is chosen so that the on-chip photodiode array can be completely illuminated, but the light does not extend beyond the surrounding on-chip light shield. Photocurrent due to background illumination is cancelled by covering the aperture and subtracting the residual photocurrent. The imager chip is then replaced with a Newport 818-SL photodetector. The aperture is increased to 1.52 mm diameter to ensure a more accurate reading. As a check, the aperture is removed to completely illuminate the photodiode.

The responsivity is

$$S = \frac{I_{pdarray}/A_{pdarray}}{P_{818-SL}/A_{818-SL}}, \tag{5.1}$$

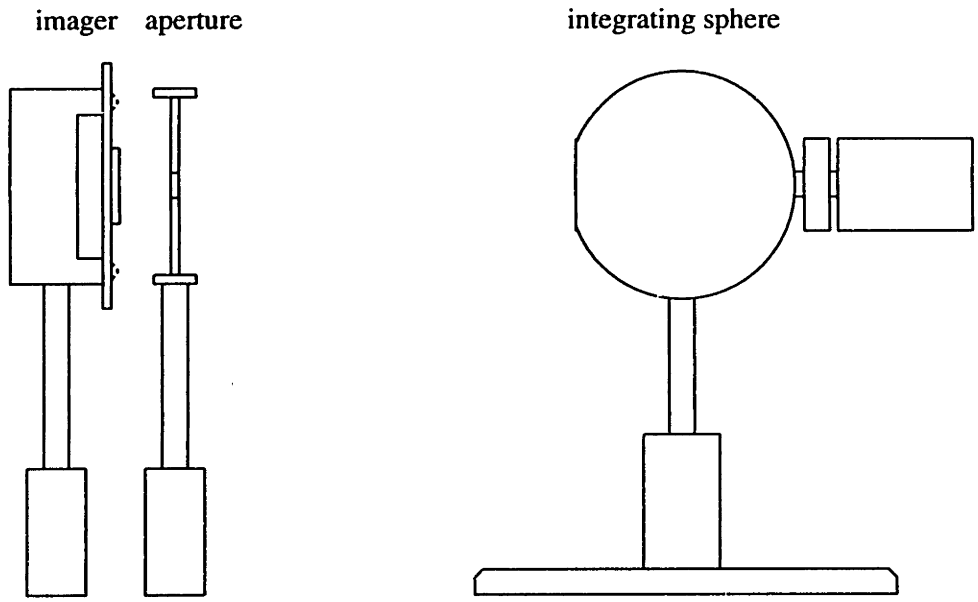


Figure 5-2: Test setup for measuring pixel responsivity.

where $I_{pdarray}$ is the photodiode array current with background illumination cancelled, $A_{pdarray}$ is the area of the on-chip photodiode array, P_{818-SL} is the power measured by the 1815-C optical power meter, and A_{818-SL} is the illuminated area of the 818-SL photodetector. The photodiode array area is

$$A_{pdarray} = 100 \cdot (22\mu\text{m})^2 = 4.84 \cdot 10^{-8}\text{m}^2. \tag{5.2}$$

The illuminated area of the 818-SL is approximately $1.8 \cdot 10^{-6}\text{m}^2$ with the aperture and $1.13 \cdot 10^{-4}\text{m}^2$ without the aperture [56]. Combining these expressions with the measured data results in the responsivities shown in Table 5-1.

The responsivities with and without the aperture should be the same, since the area is taken into account in both cases. Any difference between the two readings is due either to experimental error or a λ-dependent difference between the photodetector responsivities at the center and edge.

Conversion Gain

Conversion gain is the ratio of output voltage to integrated charge. Equivalently, it is the ratio of NMOS source follower gain to sense capacitance. Because the sense capacitance and source follower gain are nonlinear, conversion gain is not uniform over the output voltage range. An average value can be obtained by finding the charge required to move the pixel output voltage

Table 5-1: Responsivity of photodiode array.

λ (nm)	$I_{pdarray}$ (pA)	$P^{(1)}_{818-SL}$ (nW)	$P^{(2)}_{818-SL}$ (μW)	$S^{(1)}$ (mA/W)	$S^{(2)}$ (mA/W)
450	11.1	1.93	1.55	14.9	16.7
550	25.3	2.99	2.41	22.0	24.5
650	52.9	4.60	4.48	29.9	27.6

[1] with aperture

[2] without aperture

over its output range. Simulation predicts an average conversion gain of 16.8 $\mu V/e^-$. The measured value is 13.1 $\mu V/e^-$.

Compressive Transfer Characteristic

The dynamic range enhancement technique is tested by applying a staircase waveform to the lateral overflow gate. The compression curve used is quasi-logarithmic; for some range of pixel photocurrent, the relation between photocurrent and output voltage is approximately logarithmic.

$$Q(I) = \begin{cases} \frac{127}{128}TI & 0 \le I \le I^* \\ \frac{1}{7}Q_{max} + \frac{63}{128}TI & 2^0 I^* \le I \le 2^1 I^* \\ \frac{2}{7}Q_{max} + \frac{31}{128}TI & 2^1 I^* \le I \le 2^2 I^* \\ \frac{3}{7}Q_{max} + \frac{15}{128}TI & 2^2 I^* \le I \le 2^3 I^* \\ \frac{4}{7}Q_{max} + \frac{7}{128}TI & 2^3 I^* \le I \le 2^4 I^* \\ \frac{5}{7}Q_{max} + \frac{3}{128}TI & 2^4 I^* \le I \le 2^5 I^* \\ \frac{6}{7}Q_{max} + \frac{1}{128}TI & 2^5 I^* \le I \le 2^6 I^* \\ Q_{max} & I \ge 2^6 I^* \end{cases} \tag{5.3}$$

where

$$I^* = \frac{2}{7}\frac{Q_{max}}{T}. \tag{5.4}$$

The corresponding barrier waveform is

$$b(t) = \begin{cases} \frac{1}{8}Q_{max} & 0 \le t \le \frac{1}{2}T \\ \frac{2}{8}Q_{max} & \frac{1}{2}T \le t \le \frac{3}{4}T \\ \frac{3}{8}Q_{max} & \frac{3}{4}T \le t \le \frac{7}{8}T \\ \frac{4}{8}Q_{max} & \frac{7}{8}T \le t \le \frac{15}{16}T \\ \frac{5}{8}Q_{max} & \frac{15}{16}T \le t \le \frac{31}{32}T \\ \frac{6}{8}Q_{max} & \frac{31}{32}T \le t \le \frac{63}{64}T \\ \frac{7}{8}Q_{max} & \frac{63}{64}T \le t \le \frac{127}{128}T \\ 0 & \frac{127}{128}T \le t \le T \end{cases} \tag{5.5}$$

The theoretical barrier curve and compression curve are plotted in Fig. 5-3. The lowest potential applied to the lateral overflow gate must keep the current source in saturation. The minimum voltage on the current source is 315 mV. The minimum sense capacitor voltage is chosen to give a source voltage of 500 mV, and the lateral overflow gate voltage is chosen to lie 700 mV above this level. This gives a minimum lateral overflow gate voltage of 2.2 V. The maximum voltage is 5 V. The other six levels are evenly spaced between these limits. The measured and simulated data for the compression curve are shown in Fig. 5-4. Note that the pixel output voltage does not become zero in the limit of zero photocurrent; this is due to feedthrough from the barrier gate overlap capacitance.

Sample Pixel Output Waveforms

The data presented for the quasi-logarithmic compression curve shows the difference between the pixel output voltages at the end of the integration period and during reset. It is also interesting to examine the pixel output voltage as a function of time over the entire integration period. Figures 5-5, 5-6, and 5-7 show the pixel response under low, medium, and high illumination, respectively.

Two particular features are noteworthy. The low illumination figure shows significant feedthrough from the barrier gate to the pixel output at each step. Feedthrough is actually present in all cases, but is easiest to see in the low illumination case. The other feature is easiest to see in the medium illumination case. As the pixel output saturates at a barrier step, the pixel output does not abruptly change from a linear region to a flat line. The settling looks approximately exponential when the pixel output is close to saturation. This is not evident in the low illumination case because the pixel output is never close to saturation. It is not visible in the high illumination case because the photocurrent is much higher and the settling

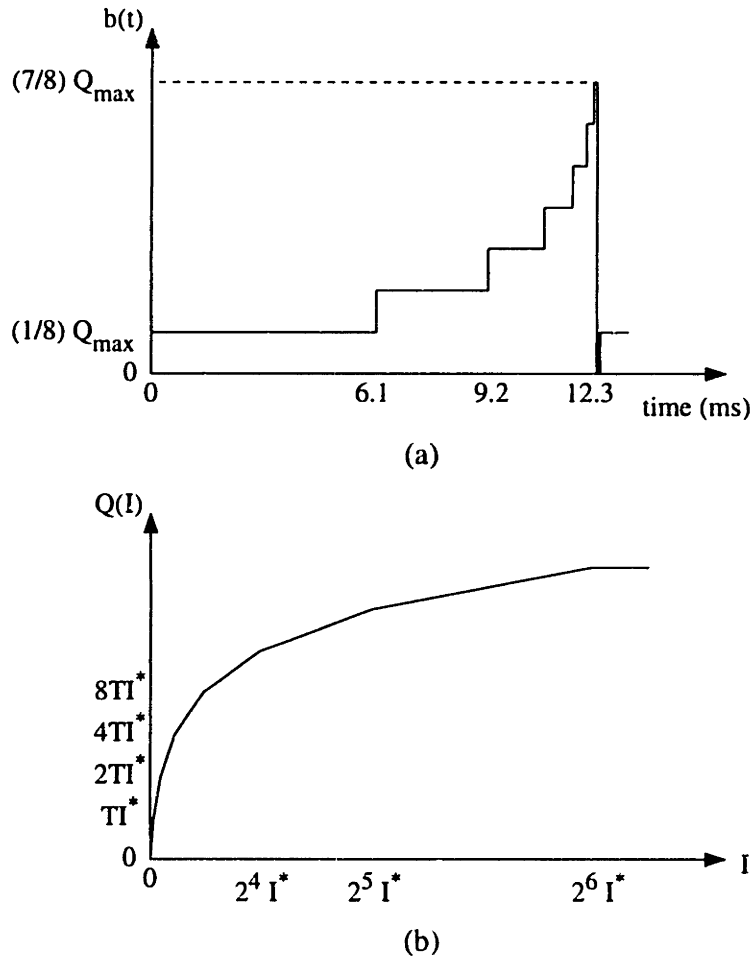


Figure 5-3: Theoretical barrier and compression curves for pixel compression test. (a) Barrier curve (b) Compression curve.

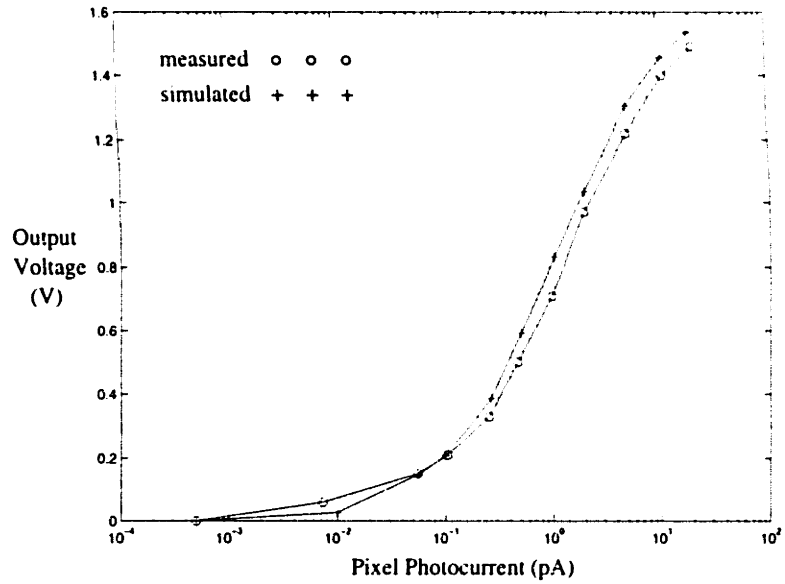


Figure 5-4: Simulated and measured pixel output using quasi-logarithmic compression curve.

is therefore much faster.

5.1.2 Analog-to-Digital Converter

The test board circuits used to test the on-chip analog-to-digital converter are shown in Fig. 5-8. The host computer sets the input by sending a 12b digital word to the DAC. The DAC output is a fully differential current. An offset current is added to both sides via a resistive divider, and the resultant current is converted to voltage by passing it through a V-to-I converter. The system gain is tuned by adjusting the bias current in the DAC. The system common-mode offset is adjusted by either R1 or R2 and R3, while the differential offset is adjusted by R2 and R3. The potentiometers are adjusted to produce a full-scale swing of 2 V and a 2 V common-mode voltage.

The output bits are available alternately at **tacpoutb** and **tacpouta** during the **phi1** and **phi3** phases, respectively. During its off phase, an ou put bit is zero. Negative logic is used; a low output represents 1 and a high output represents ο The NAND gate properly combines information from both outputs so that the bits shifted into the shift register use positive logic. The input to the shift register is latched on the rising edge of the clock, which is delayed by 25 ns relative to the data to allow for propagation delay. The timing is shown in Fig. 5-9. After 12 bits have been shifted into the shift register, the contents are transfered to a register. The

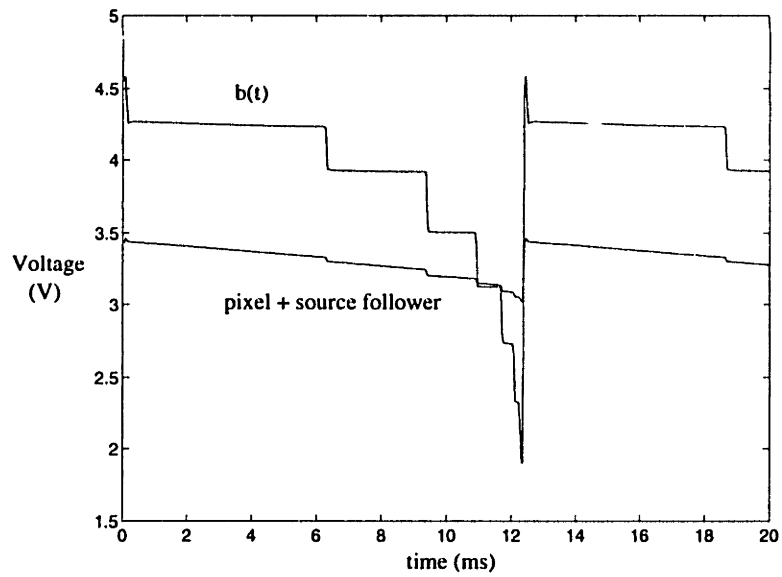


Figure 5-5: Measured pixel output and barrier curve for low illumination.

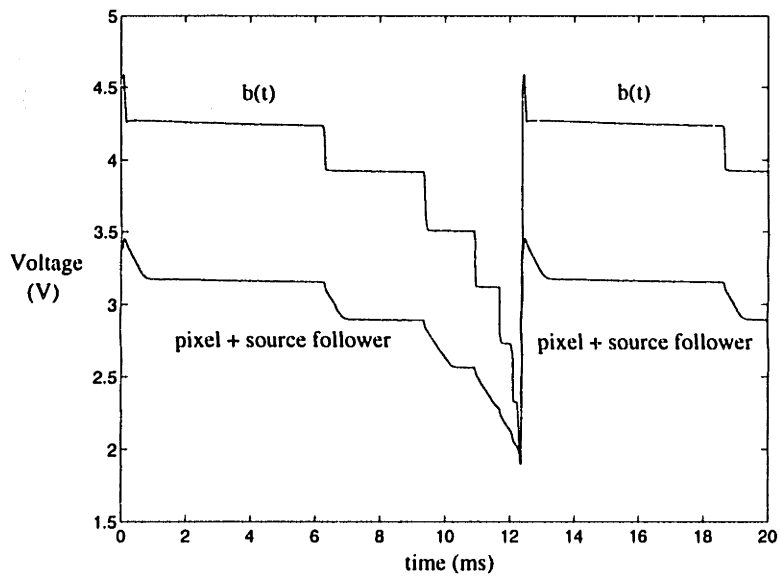


Figure 5-6: Measured pixel output and barrier curve for medium illumination.

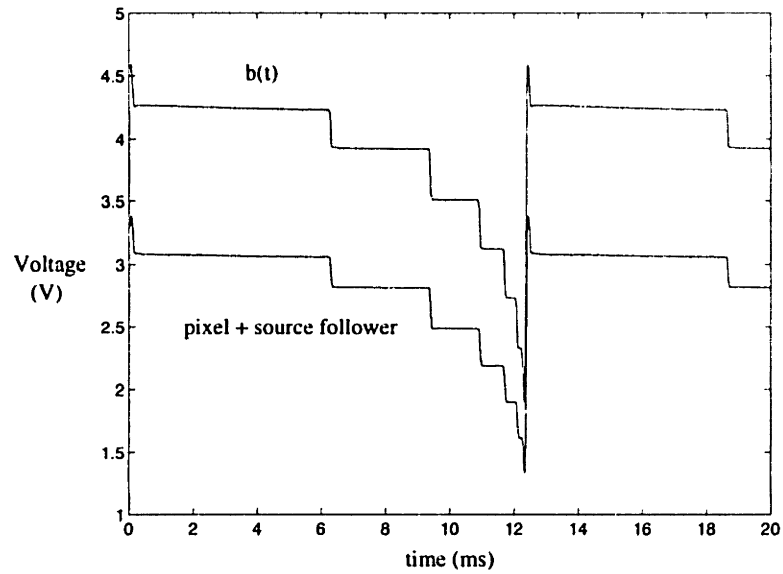


Figure 5-7: Measured pixel output and barrier curve for high illumination.

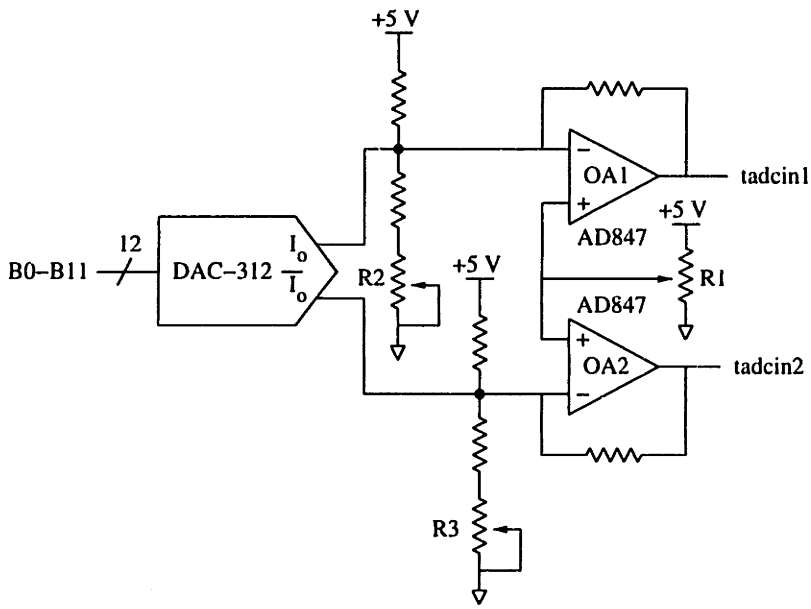
host computer reads this value, increments the digital word sent to the DAC, and repeats the procedure until all input values are tested.

The transfer characteristic for the on-chip ADC is shown in Fig. 5-10. The transfer characteristic was obtained by averaging the results of 100 sweeps. The tested sampling rate is 12.8 kS/s, and the sample is digitized to 12 bits. The differential and integral nonlinearity are shown in Fig. 5-11 and Fig. 5-12, respectively. The DNL and INL indicate 9-b accuracy. The INL plot has a large shift at the major carry point and smaller shifts at the next most significant carry points, characteristic of capacitor mismatch or signal-dependent charge feedthrough. Nonlinearity plots obtained from three chips produced very similar results.

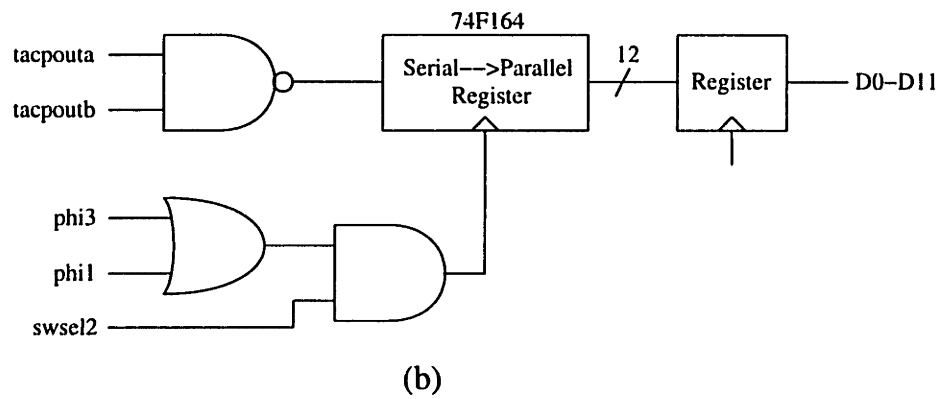
5.1.3 Correlated Double-Sampling Circuit

The on-chip correlated double-sampling circuit is tested by applying a stepped voltage at the input and observing the output voltage. The important parameters are offset and gain. Offset is the output differential voltage given a 0 V step input, and slope is the ratio of output differential voltage to input step voltage. Linearity should be very good, but is not a critical parameter since the pixel output itself is highly nonlinear.

The test board circuits for generating the input test signal and sensing the CDS output are shown in Fig. 5-13(a) and (b), respectively. The host computer generates a 12b word, which generates the analog voltage **tadcin1** in Fig. 5-8. Signal **tadcin1** simulates the pixel output



(a)



(b)

Figure 5-8: (a) Test board circuit to generate test input signal for on-chip ADC. (b) Test board output data path for on-chip ADC.

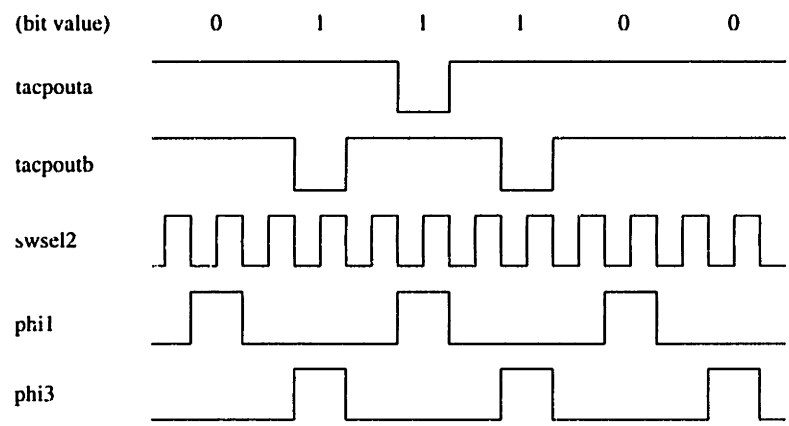


Figure 5-9: Timing for test analog-to-digital converter.

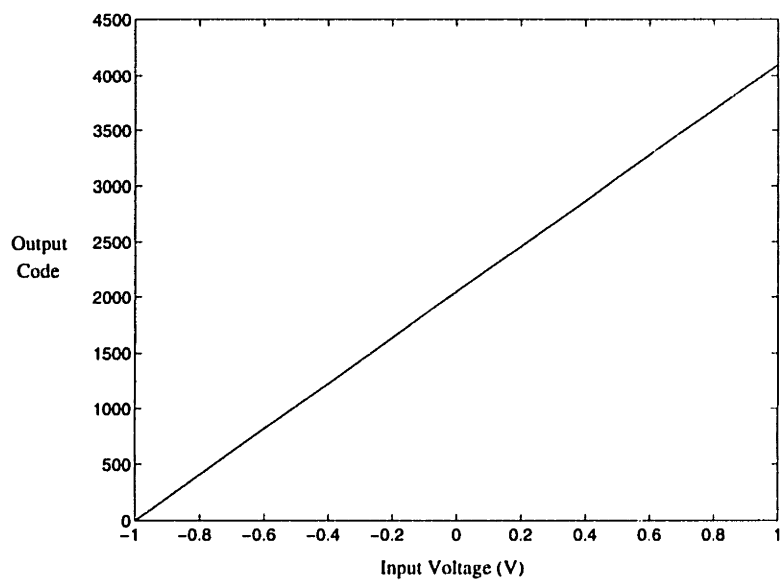


Figure 5-10: Transfer characteristic for on-chip analog-to-digital converter.

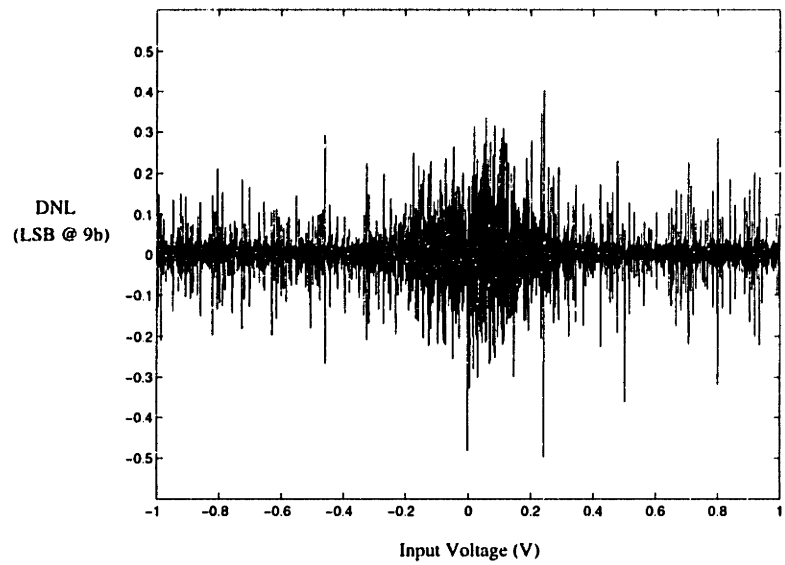


Figure 5-11: Differential nonlinearity plot for on-chip analog-to-digital converter.

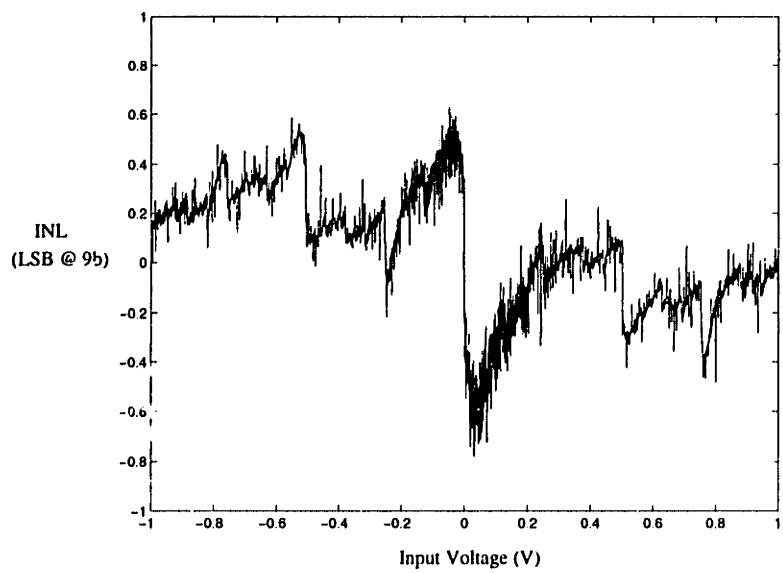


Figure 5-12: Integral nonlinearity plot for on-chip analog-to-digital converter.

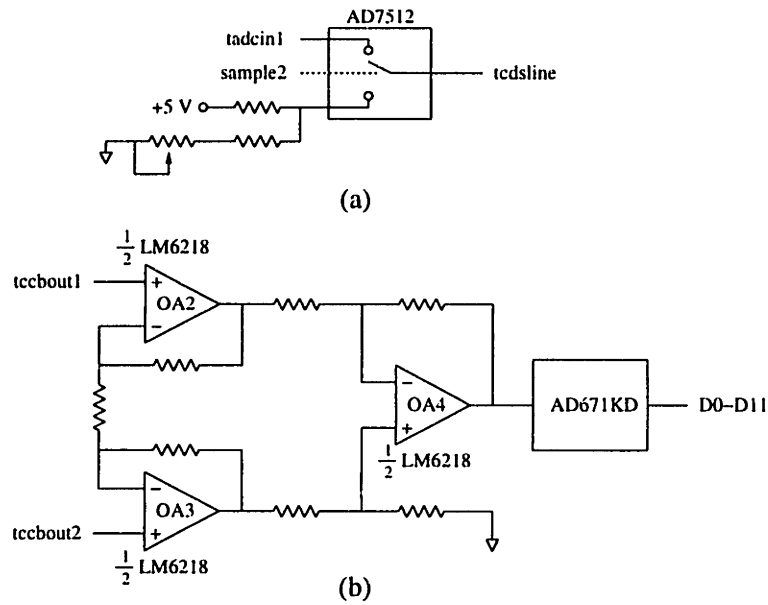


Figure 5-13: (a) Test board circuit to generate input signal for CDS circuit (b) Test board circuit for sensing CDS output voltage.

voltage at the end of the integration period. The analog multiplexer (AD7512) switches between this voltage and a constant voltage simulating the pixel output voltage at reset. The stepped voltage, **tcdsline**, is input to the on-chip CDS circuit.

The on-chip CDS circuit outputs are **tccbout1** and **tccbout2**. Test board amplifiers OA2, OA3, and OA4 implement an instrumentation amplifier, performing a differential to single-ended conversion on **tccbout1** and **tccbout2**. The single-ended signal is converted by an AD671KD analog-to-digital converter. The output bits are latched into a register and down-loaded to the host computer. The host computer increments the digital input word and repeats the procedure until all input values are tested.

Testing of the CDS circuit verified that the output transfer characteristic (V_{out} vs. input step voltage) was a straight line with unity slope. The output voltage is zero with a 1 V step input.

5.2 Image Frames

Frame data is used to measure the saturation level, optical crosstalk, linear and compressive imaging transfer functions, FPN, random noise, and power dissipation. The fundamental clock period for most tests is 814 ns, giving an integration time of 40 ms and an equivalent frame

rate of 30 frames/sec at 10 bits/sample.

5.2.1 Saturation Level

Saturation is obtained by subtracting the average pixel value in a dark frame from the average pixel value in a saturated frame, using linear imaging. The difference in average pixel value is 3451 codes, corresponding to 1.69 V.

5.2.2 Dark Response

Dark response is measured by keeping the imager in the dark and calculating the average pixel value for several different integration times. The longest integration time used is 1.28 sec, and the shortest is 20 ms. The measured dark response is 174 mV/s, which is equivalant to a pixel dark current of approximately 2.1 fA.

5.2.3 Optical Crosstalk/Blooming

Blooming and optical crosstalk are tested by applying high illumination to a small region of the imager. Blooming and optical crosstalk appear as a saturated region much larger than the area actually exposed to illumination. An informal test was performed by imaging a tubular fluorescent lamp in front of a newspaper. The acquired image shown in Fig. 5-14 shows no sign of blooming.

In another test, a laser beam was directed onto the imager. A 0.4 mW HeNe laser with a wavelength of approximately 650 nm was used, and the illumination was varied using neutral density filters. The glass cover on the imager chip was removed to avoid any scattering or reflection from that surface. The benefit of the laser beam is that it minimizes the effect of any reflective paths. The direct beam falls completely on the pixel array. Much of the incident beam is reflected off of the imager, but as long as there are no specular surface nearby, any reflected light that returns to the imager is greatly attenuated. The results are shown in Fig. 5-15. At the higher illuminations, the central spot is surrounded by an interesting pattern of light spots. These spots are not characteristic of blooming or optical crosstalk, since these would create a (roughly) circularly symmetric spot with a diameter greater than the exposed spot. The spots are apparently due to the laser itself, since they can be observed in the incident beam by holding a sheet of paper in front of the laser. The conclusion is that blooming does not occur below an illumination of $\sim 50,000\times$ saturation illumination in the linear mode.

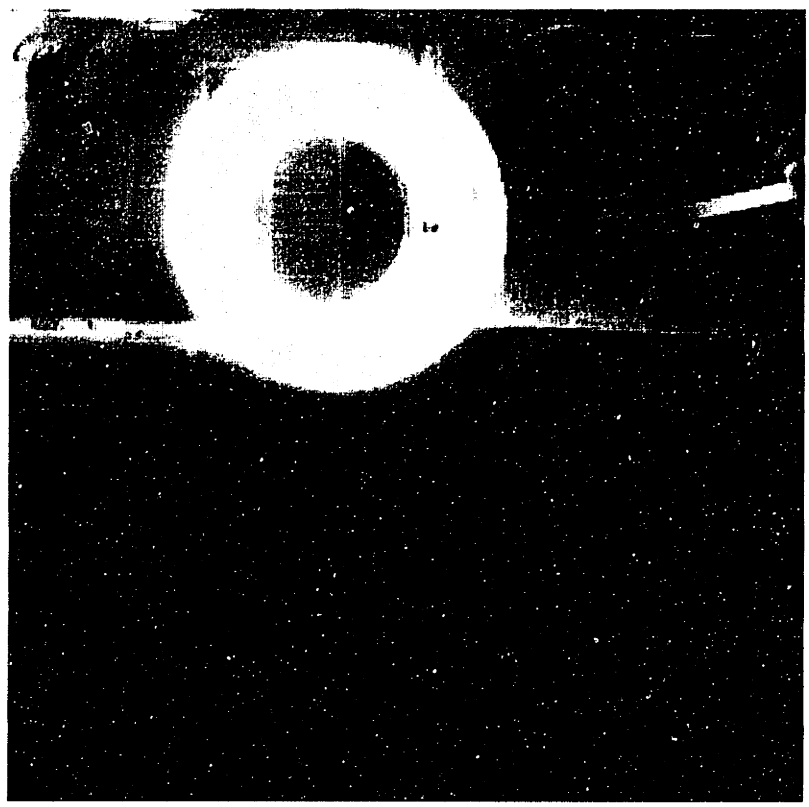


Figure 5-14: Blooming test performed with fluorescent lamp and newspaper. The horizontal bright line is a crease in the newspaper.

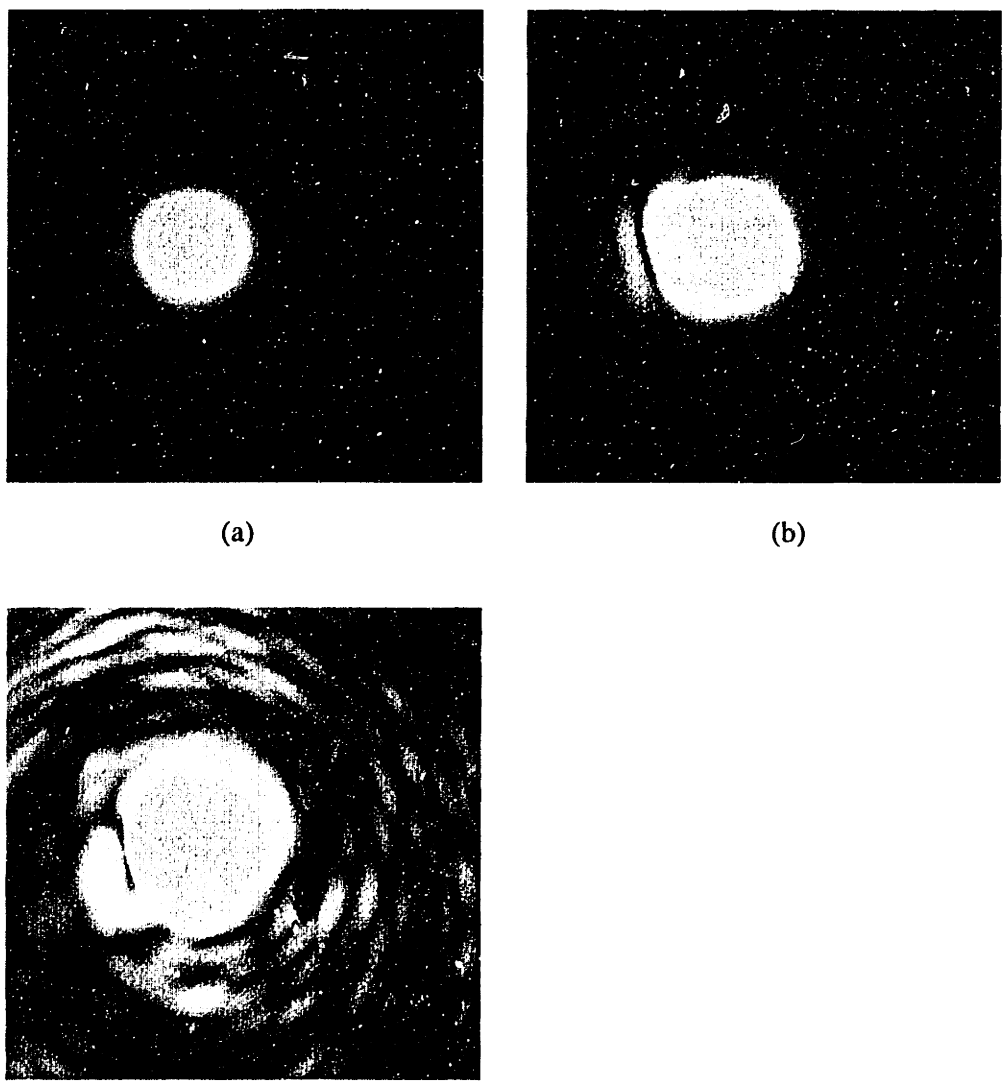


Figure 5-15: Blooming test performed with laser. Linear imaging mode used in all cases. (a) $\sim 500\times$ saturation illumination. (b) $\sim 5,000\times$ saturation illumination. (c) $\sim 50,000\times$ saturation illumination.

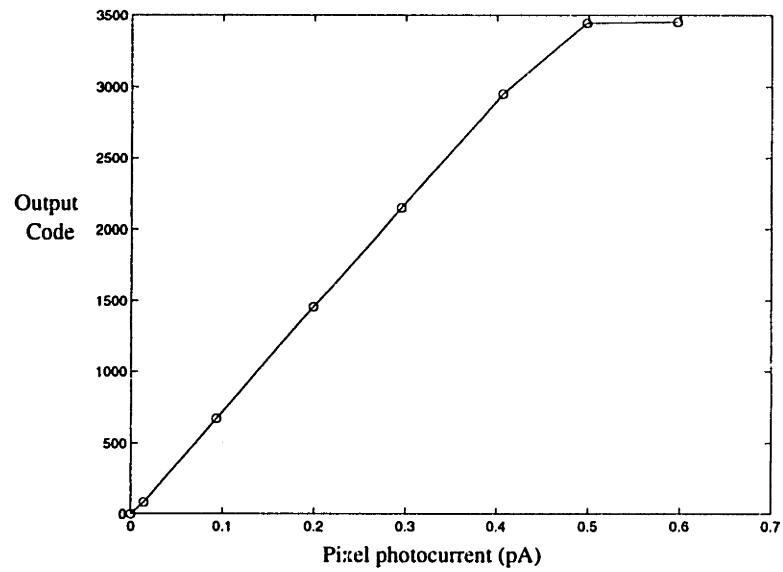


Figure 5-16: Measured transfer characteristic for linear imaging mode.

5.2.4 Transfer Characteristic

Transfer characteristics were measured for the imager operating in the linear mode and the compressive mode. The linear transfer characteristic is shown in Fig. 5-16. The compressive transfer characteristic is shown in Figs. 5-17 and 5-18, along with simulated results from HSPICE. The barrier curve used to obtain the compressive transfer function is shown in Fig. 5-19. Figure 5-20 shows the transfer characteristic for a barrier curve intended to produce a high dynamic range; the measured dynamic range increase is $\sim 160\times$.

5.2.5 Fixed-Pattern and Random Noise

Fixed-pattern noise (FPN) and random noise are determined by uniformly illuminating the pixel array at several different irradiances. At each irradiance, 16 frames are captured. FPN is calculated from the measured data using Eqs. (2.1) and (2.2). Random noise is determined by averaging all 16 frames, and subtracting that average from each frame. The rms pixel value of the difference frames represents random noise.

$$v_{n,rms} = \sqrt{\frac{\sum_f (P_f(i,j) - \overline{P}(i,j))^2}{f-1}}. \tag{5.6}$$

The FPN and random noise for the linear imaging mode is shown in Fig. 5-21. Pixel FPN initially rises from its value in the dark due to gain FPN sources, but falls when the pixels

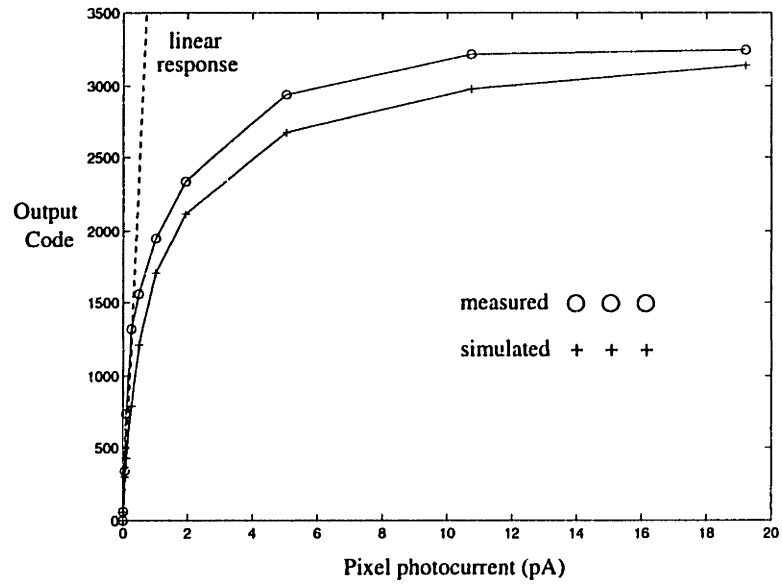


Figure 5-17: Measured and simulated transfer characteristics for compressive imaging mode (linear plot). Linear response transfer function is shown for comparison.

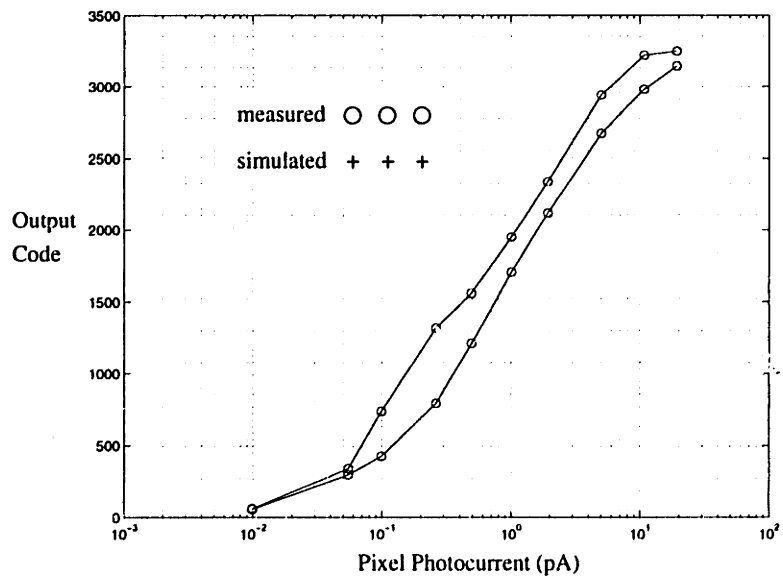


Figure 5-18: Measured and simulated transfer characteristics for compressive imaging mode (log plot).

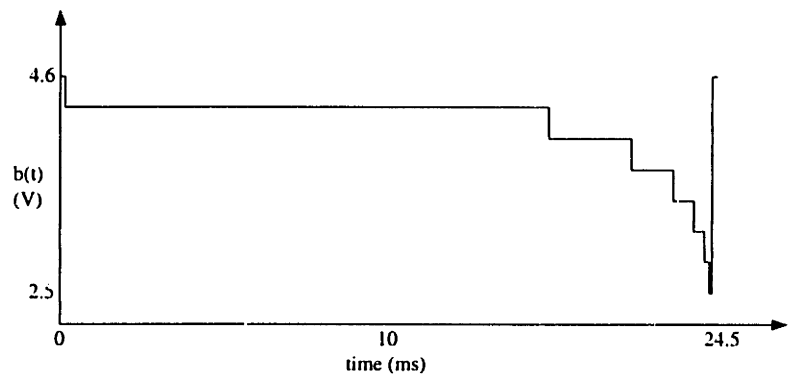


Figure 5-19: Barrier curve for compressive imaging mode.

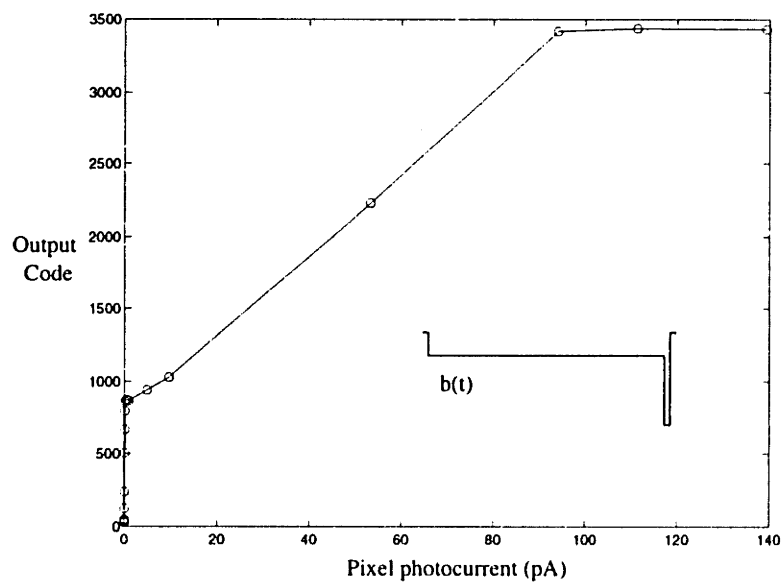


Figure 5-20: Compression curve and barrier curve (inset) for high dynamic range expansion.

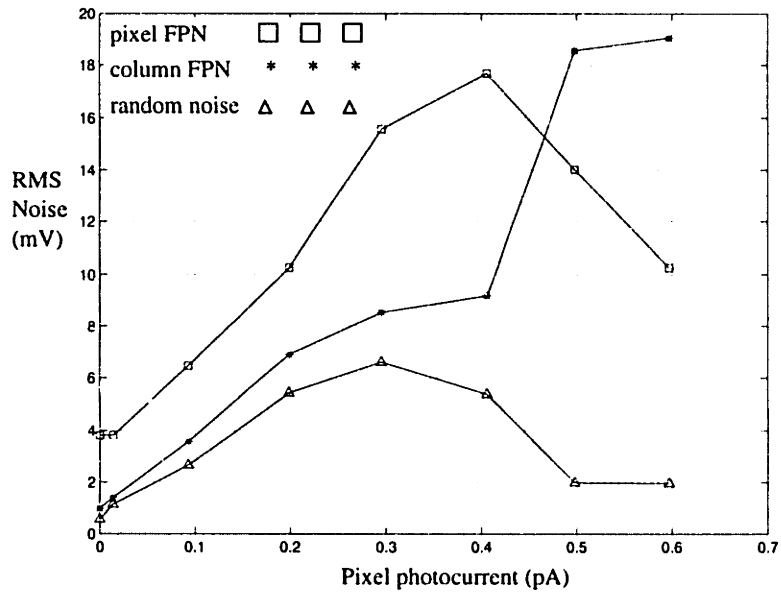


Figure 5-21: Measured random and fixed-pattern noise for imager operating in linear mode.

saturate. This is expected, since sense capacitance and optical aperture do not affect the output voltage when the pixel saturates. Random noise also initially increases from its value in the dark, presumably due to photonic noise. It falls when the pixel saturates, since photonic shot noise goes away, leaving only barrier shot noise.

The FPN and random noise for the compressive imaging mode is shown in Fig. 5-21. Compared with the FPN and random noise for the linear imaging mode, random noise and pixel FPN are clearly reduced. This is due to the limited free charge Q_{free}, which is the sole source of photonic noise and gain FPN. Column FPN approximately follows the shape of the transfer characteristic, implying that column-to-column mismatch is mostly a gain error on the pixel output voltage.

5.2.6 Sample Frames

Figure 5-23 is a frame sequence which demonstrate the wide dynamic range feature. In Fig. 5-23(a), the imager operates in the linear imaging mode. The illumination is sufficient for the upper half of the truck, but the lower half is underexposed. In Fig. 5-23(b), the overall illumination has been increased so that detail in the lower half of the truck is visible, but now the upper half of the truck is overexposed. In Fig. 5-23(c), the imager operates in the compressive imaging mode, using the same illumination as in (b). Details in both halves of the truck

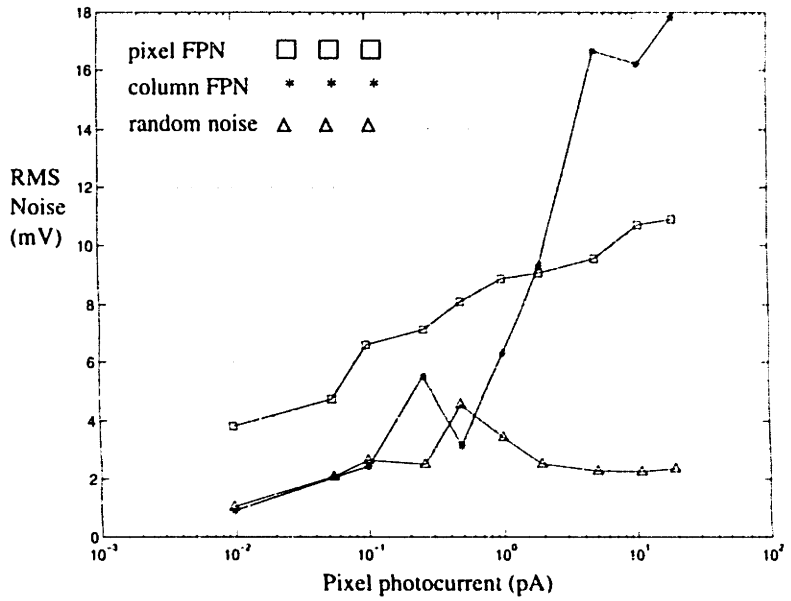


Figure 5-22: Measured random and fixed-pattern noise for imager operating in compressive imaging mode.

are visible, demonstrating the value of the wide dynamic range feature.

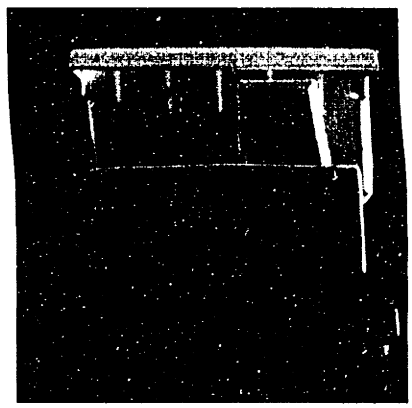
5.2.7 Imager Power Dissipation

Power dissipation was determined by monitoring the difference in the supply current to the board with and without the imager inserted. The on-chip digital power dissipation, except for the pad driver power, could not be determined due to large fluctuations in the measured current. This power can be roughly estimated by examining all nodes which are charged through the digital supply. For each such node, the energy dissipated per frame is

$$E_{diss} = \frac{1}{2} C V^2 N_{sw}, \tag{5.7}$$

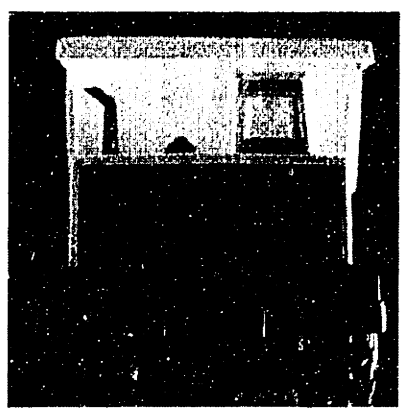
where C is the node capacitance, V is the supply voltage (5 V), and N_{sw} is the expected number of times that the signal changes state over one frame. In order to simplify the calculation, C includes only gate and diffusion capacitance; wiring capacitance is ignored. Also, the short-circuit current during switching is ignored. Summing all of these terms for the entire imager, the energy is estimated to be 16.5 μJ/frame. At a 30 Hz frame rate, this is only about 0.5 mW, and thus is quite small compared to the other components.

Power dissipation as a function of frame rate is shown in Fig. 5-24. At each measured frame rate, power was determined by decreasing the bias currents until the image became visibly



(a)

(b)



(c)

Figure 5-23: Sample frames demonstrating wide dynamic range feature. (a) Linear imaging with underexposure of bottom half. (b) Linear imaging with overexposure of top half. (c) Wide dynamic range imaging using same illumination as (b).

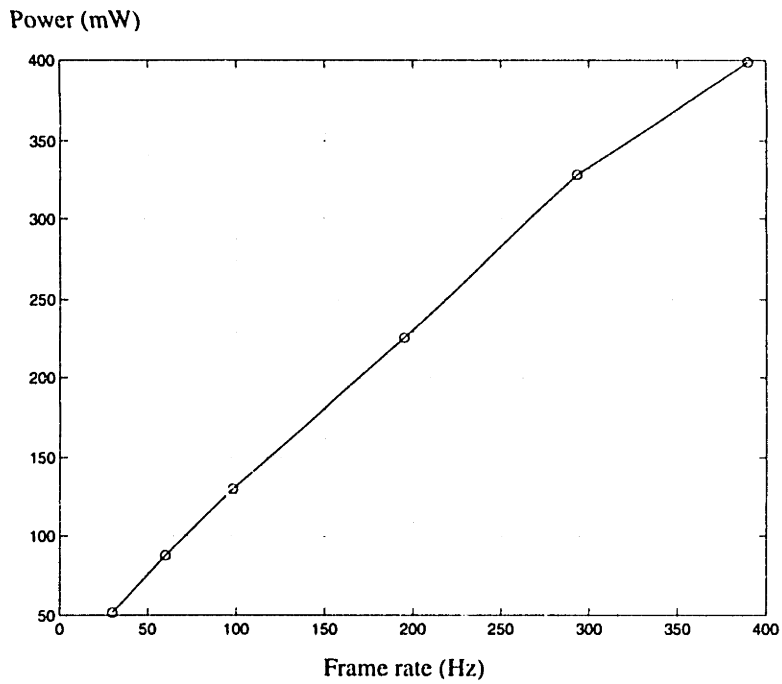


Figure 5-24: Frame rate vs. imager power dissipation (5 V). Power quoted is sum of analog supply power and output pad driver power.

degraded. Assuming that the curve is linear out to 1000 fps, it would take approximately 1 W to operate the imager at 1000 fps.

5.2.8 Performance Summary

The measured and estimated imager performance is summarized in Table 5-2. With the exception of responsivity, measured and estimated figures are in close agreement. Dark response was not estimated, since dark current is extremely process dependent. The conversion gain was estimated by HSPICE simulation using BSIM level 28 parameters with a pixel circuit model extracted from the layout. Fixed-pattern noise is estimated from Fig. 3.6; lacking actual mismatch data for the process, the $\sim 2\times$ difference between measured and estimated value is reasonable. Estimated saturation level was obtained from an HSPICE simulation of the pixel. Random noise is estimated from dark shot noise and reset (kT/C_{sense}) noise. The measured dark current and sense capacitance were used in the random noise estimate. The measured noise performance is actually better than estimated. The reason may be that the barrier noises during the sample and reset phases are correlated due to the slow time constant at the sense node.

Table 5-2: Summary of imager performance.

parameter	measured	estimated
dark response	174 mV/s @ 22° C	—
responsivity	23 mA/W @ 550 nm	82 mA/W @ 550 nm
conversion gain	13.1 μV/e$^-$	16.8 μV/e$^-$
FPN (dark)	4.0 mV (1 σ)	2.3 mV (1 σ)
power dissipation	52 mW @ 30 Hz, 5 V (10 bits)	—
saturation level	1.69 V	1.64 V
random noise (dark)	0.56 mV (1 σ)	0.88 mV (1 σ)
DR (linear mode)	3000	2000
DR expansion	1× - 160×	$\sim$ 200×

The FPN in the linear mode consists of an offset term of 4 mV rms (0.2% saturation) and a gain term of approximately 1% of the output signal. The offset and gain FPN are comparable to other recent CMOS imagers, where such comparison can be made. For example, a series of 128 × 128 APS arrays achieved p-p dark FPN ranging from 3.7% saturation to 0.08% saturation, depending on array implementation and whether or not CDS was used [58]. If the peak-to-peak value corresponds to the 3 σ level, for example, then the rms FPN's would range from 1.2% saturation to 0.03% saturation. Another recent 256 × 256 APS array reports dark FPN of 0.2% saturation p-p, and FPN arising from conversion gain mismatch as "under 1%" [57]. The corresponding dark FPN value is 0.07% saturation rms. A 128 × 128 APS array with electronic shutter quoted a dark intracolumn (pixel) FPN of 1.5 mV and a bright intracolumn FPN of 5.1 mV, which represents a change of 3.6 mV on a saturation level of 470 mV, or 0.76% gain FPN rms[59]. The dark pixel FPN is 0.3% saturation. On the same imager, column FPN increased from 12.0 mV to 18.6 mV rms between the dark and bright frames, representing about 1.4% gain FPN.

Chapter 6

Conclusions

6.1 Summary

This thesis described the design and testing of a CMOS imager for automotive machine vision. The goal was a high speed, high dynamic range imager which allows a high level of system integration.

On-chip column-parallel analog-to-digital conversion was chosen for several reasons. Column-parallel outputs allow high frame rates. Digital output allows convenient integration with digital signal processors, without crosstalk between adjacent channels. Integrating the converters on the imager chip potentially reduces system size, power, and cost.

Several converter architectures were compared for focal-plane conversion. Full flash converters are too large for the required 10 bit accuracy. Dual-slope converters are usable for standard frame rate (30 frames/sec), but too slow for 1000 frames/sec operation. Delta-sigma converters are also too slow, and very poorly matched to the signal format coming from a pixel array. Algorithmic converters do not suffer from these problems. Cyclic and pipeline algorithmic converters are both suitable choices; a cyclic architecture was chosen.

Wide dynamic range is required due to uncontrolled lighting conditions in automotive applications. Several previous schemes for increasing dynamic range were reviewed. These schemes could be roughly divided into two types: nonlinear pixels and multiple integration periods. Nonlinear pixels usually use the logarithmic current-to-voltage characteristic provided by a forward-biased diode or a subthreshold MOSFET. These techniques provide very large dynamic range, but no control over the compressive characteristic. Multiple integration periods are more flexible, but often require frame storage which makes integration difficult. Two CCD-

based schemes were described which do not require frame storage. For one of these schemes, the dynamic range increase is set by a gate area ratio, which is severely limits the achievable improvement. The other is based on time period ratios, which is much more useful.

The wide dynamic range pixel employed in this imager was essentially identical to a CCD version first described by Knight. It was shown that a CMOS pixel could employ the same technique, and that a stepped barrier function could be employed with little loss of flexibility. The barrier is digitally controlled from off-chip by just three digital signals. A 256-stage digital shift register selects one of eight barrier levels for each row. This shift register, and a row select shift register, are coordinated with the correlated double-sampling circuits to read out the array in "moving shutter" fashion. This avoids the need for storage in the array, which reduces the influence of optical crosstalk and increases fill factor.

A prototype imaging chip was designed and fabricated in a 0.8 μm CMOS process, which is essentially a digital process with linear capacitors. The pixel is an active pixel sensor (APS) design with a charge-spill gate to increase sensitivity. The correlated double-sampling (CDS) circuit reduces FPN, but does not entirely remove it. In particular, it does not affect FPN produced by barrier gate overlap capacitance mismatch. The CDS circuit is switched-capacitor and fully-differential. Cyclic ADC's are shared between adjacent columns. Each switched-capacitor ADC is fully-differential for improved noise and mismatch immunity. Digital output is provided through 32 channels, which must be multiplexed between the 128 converters.

The pixel was characterized for dark response, responsivity, saturation level, and conversion gain. The dark response was determined by optically shielding the array and measuring the average output voltage for different integration periods; the measured dark response is 174 mV/s. The responsivity is found from illuminating an on-chip photodiode array with monochromatic light, and comparing the measured photocurrent with the reading from a power meter exposed to the same illumination; the result is 9.8 mA/W at a 550 nm wavelength. The saturation level is determined from the minimum and maximum output codes from the array; the measured result is 1.69 V. Conversion gain is determined from the change in average output code when light intensity is varied; the conversion gain is 13.1 μV/ e^-.

The CDS and ADC transfer characteristics were obtained from separate test structures on the chip; the outputs were found to be well-behaved. The ADC was further characterized for nonlinearity. A set of 100 sweeps were performed and averaged to obtain the DNL and INL; they were determined to be compatible with 9 bit performance. Their accuracy is limited by capacitor mismatch, which was expected.

FPN was found by taking sets of 16 frames at each of multiple illumination levels. The

frames at each illumination were averaged to remove random noise. The random noise in the dark was 4.0 mV (1 σ). Random noise was found by subtracting the FPN from each of the 16 frames. The random noise in the dark was 0.56 mV, which provides a base dynamic range for the linear imager of 3000:1.

Power as a function of speed was determined by setting the bias currents at each speed to the minimum required to remove any degradation in the image. The power dissipation was approximately linear with speed, increasing from 52 mW at 30 fps to about 400 mW at 400 fps. It was actually possible to obtain good quality images at 800 fps, but the power could not be increased enough to completely remove all degradation. Power was obtained by adding the contributions of the analog supply power and the power for the output pad drivers. It was not practical to measure the digital supply power, but this component should be quite low at low to moderate frame rates.

The wide dynamic range technique was verified, both at the pixel level and at the frame level. Transfer characteristics were obtained for quasi-logarithmic compression. These were found to agree with simulated results. Frame dynamic range was investigated by illuminating the imager with a laser beam. No blooming was found for illuminations as high as 50,000× saturation exposure (in the linear mode).

Several factors limit the usable dynamic range. One limitation is internal reflections in the camera body. Reflections can occur between the lens, imager chip surface, and the lens mount. The camera board limits dynamic range by imposing a minimum integration time of 1/256 of the frame time. The major limitation of the imager iteself is the minimum time required to sample a pixel output voltage, which depends on how fast the CDS circuit can settle. Since this can settle in about 1 μs, the maximum dynamic range in this case is on the order of 30,000 for a 30 ms integration time.

A camera and software were developed. The camera consists of a camera board, lens and lens mount, a power supply, and a host PC. The camera board provides bias voltages, bias currents, and clocking. Output frames are stored in on-board memory; approximately 2.5 frames of full 12-bit data can be stored. The host PC downloads frame data through a bidirectional parallel port.

6.2 Contributions

The main contribution of this thesis is the integration of the wide dynamic range pixel into an APS architecture. The starting point of the thesis was a scheme for applying a continuous-time

barrier function to a lateral overflow gate in a single pixel. This work has demonstrated a practical imager using these pixels with the following properties:

- No charge storage required in the pixel or a vertical shift register, which means higher fill factor and higher achievable dynamic range.
- Digital generation and control of the barrier function. Only n control bits are required to produce a 2^n-slope barrier function.
- Offset introduced by threshold voltage mismatch in the barrier device is partially removed by correlated double-sampling.

A secondary contribution is the switched-capacitor correlated double-sampling circuit. The CDS circuit in this thesis simultaneously removes offsets and converts the single-ended pixel output voltage to a fully-differential voltage. Charge injection from the switches is cancelled, to first order, by the fully-differential design of the CDS.

A third contribution is a successful demonstration of column-parallel cyclic analog-to-digital conversion for focal-plane conversion. Crosstalk is necessarily a major concern when ADC's and pixel arrays are integrated, but it is very difficult to accurately estimate. This chip is "proof-by-existence" that cyclic converters and an APS can be integrated with no apparent ill effects. even at high frame rate.

6.3 Conclusions

The original goals for the imager were

- a 64×64 pixel array
- 1000 fps operation
- 10-bit output
- fill factor $\geq 50\%$
- dynamic range of at least 10^6:1 (120 dB)

The performance of the fabricated imager in most cases approaches or exceeds these specifications.

- By switching to a 0.8 μm CMOS process, a 256×256 pixel array was achieved.

- The imager operates with no visible image degradation up to 390 fps, and with only minor degradation up to approximately 800 fps. The frame rate is limited by the relatively large capacitors used in the CDS and ADC circuits. The op-amp and associated bias circuits were originally designed assuming 50 fF unit capacitors would be used. A second lot was fabricated where the 50 fF capacitors were replaced with 400 fF capacitors, but the chip was otherwise identical. After metal mask changes to correct errors on the first run of chips, only chips from the second lot were shipped.

- The converter accuracy is limited to 9 bits, so the overall output accuracy does not meet the 10-bit goal. This is not surprising, since it was not expected that capacitor matching would be much better than 9 bits. However, the pixel charge capacity would have to be at least 10^6 e^- to achieve a signal-to-noise ratio of 2^{10}:1. Given the actual pixel capacity, the signal-to-noise ratio is limited to 380 while operating in the linear mode.

- The fill factor was 49%. The effective optical aperture is somewhat greater due to the unshielded field regions between metal lines.

- The dynamic range expansion is board-limited to approximately 200×, producing an overall dynamic range of 116 dB. This is quite close to 120 dB. The dynamic range could be increased beyond 120 dB without any change in the imager itself.

The ultimate dynamic range achievable by the wide dynamic range technique is limited by the settling time of the CDS circuit. Dynamic range is maximized by minimizing the duration of the final step in the barrier curve. This last step begins when the barrier voltage falls to its minimum value, and ends when the CDS circuit has sampled the pixel output (**sample1** goes low). Ideally, the op-amp outputs would would exactly track the pixel output voltage while **sample1** is high, and hold the last value when **sample1** goes low. In reality, the pixel output voltage is a ramp, and the voltage sampled by the op-amp has the general form

$$v_{sample} = At + Be^{-t/\tau} + C, \tag{6.1}$$

where τ is the closed-loop time constant of the op-amp. The desired response is the linear term. The exponential term determines how fast the system settles to the ramp, and therefore how long **sample1** must stay high. The constant term is due to the finite transconductance of the op-amp. While the pixel output is being sampled on the bottom capacitor plates, current must be supplied to the top plates. This requires a finite input differential voltage at the op-amp input terminals. For higher illuminations, the ramp is steeper and the constant term increases.

This results in an offset between the actual and ideal responses, even if sufficient time is allowed for settling.

The imager is suitable for its intended purpose, but its suitability for other purposes is limited. The wide dynamic range algorithm imposes lower limits on both FPN and random noise performance. FPN is produced by mismatch in the overlap capacitance of the lateral overflow gate. Random noise is produced by charge flow under the lateral overflow gate. CDS removes neither of these. CDS does remove FPN due to mismatch in the in the lateral overflow gate threshold voltage, but this requires a "soft" reset which results in image lag.

The column-parallel cyclic converter structure was a success. Despite the large number of converters, image quality was not visibly degraded and power consumption was modest.

6.4 Future Work

The next step is combination of the imager with an analog or digital processor to create a complete stereo vision system. A prudent first step is behavioral simulation of the stereo vision algorithm with simulated frame sequences containing fixed-pattern noise. This will show that the algorithm is robust against FPN, or show the need for further algorithm development.

The stereo vision system requires real-time imager operation, which is not currently possible because of the long time required to download and reorganize data. The raw data could be reorganized and reformatted in real-time using dedicated digital hardware. One method is sketched in Fig. 6-1. The bits corresponding to a single sample become available over several cycles of operation. As a partially constructed sample is read out of memory, it is combined with present data from the imager and written back to the memory. When a row of samples is fully constructed, a different area of memory is used to construct the next row of samples and the reconstructed samples are read out.

The current version of the imager chip can be used as-is in the stereo vision system. If an advanced prototype system is required, it may be worthwhile to refabricate the chip. The imager interface can be dramatically simplified by a few chip modifications. The digital output bits can be reorganized on-chip so that all bits for a given pixel are output simultaneously. Timing and bias generation can also be integrated on-chip so that the only external connections are power, ground, digital barrier control, an output digital signal to indicate the start of a frame, and the fundamental clock.

If very high dynamic range is required, it will almost certainly become necessary to reduce reflections inside the camera body. Special materials and /or coatings may be required on the

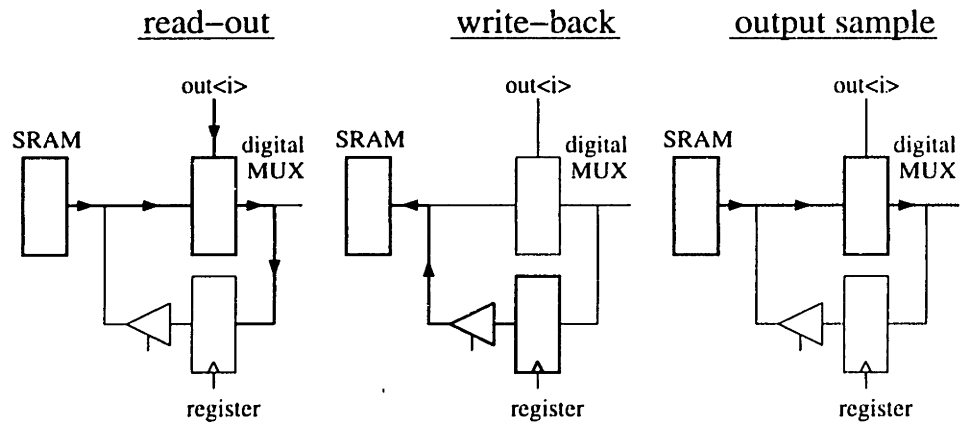


Figure 6-1: Implementation of real-time data reformatter. In the first phase, raw data from imager is combined with stored data in SRAM's. In the second phase, this data is read back into memory. In the third phase, fully reconstructed data in memory is read out as raster scan.

imager chip, lens, and lens mount.

Most commercial interest will be for consumer imagers. The correlated double-sampling and ADC circuits could be very useful for these applications, but the wide dynamic range feature is of limited use. FPN performance is very important for images that people will view, so FPN and random noise (which will appear as FPN in a digital still camera) must be kept to a minimum. As explained previously, the wide dynamic range technique sets a lower limit to random noise and FPN performance. The wide dynamic range technique extends the upper limit of usable illumination, but extension of the lower limit would be more useful.

Appendix A

Board Schematics

Abel source files for PAL1, PAL2, PAL3, PAL5, and PAL5PP are listed below (PAL4 is not used). PAL5 is used with the Strawberry Tree data acquisition card, while PAL5PP is used with the parallel port interface.

Schematic diagrams for the camera and test boards are powerview documents located in **/homes/decker/powerview/imager/camera.xxx**. They are reproduced in this appendix. The design is divided into a camera board (first three sheets) and a test board (last two sheets).

```
// pal1.abl
//
// SCCS id: 1.1
//
// Steven J. Decker
// decker@mtl.mit.edu
//
// 5/6/1997

Module pal1

Title
'PAL1 equations
Steven J. Decker
decker@mtl.mit.edu
pal1.abl
SCCS id: 1.1
5/6/97'

Declarations

pal1 device 'P22V10C';

pal_clk pin 2;
a0 pin 3;
a1 pin 4;
a2 pin 5;
a3 pin 6;
a4 pin 7;
a5 pin 9;
a6 pin 10;
a7 pin 11;

isolate pin 17 istype 'reg_D,buffer';
!sample2 pin 18 istype 'reg_D,invert';
sample1 pin 19 istype 'reg_D,buffer';
bselect pin 20 istype 'reg_D,buffer';
aselect pin 21 istype 'reg_D,buffer';
phi0 pin 23 istype 'reg_D,buffer';
!phi4 pin 24 istype 'reg_D,invert';
!phi3 pin 25 istype 'reg_D,invert';
!phi2 pin 26 istype 'reg_D,invert';
```

```
!phi1 pin 27 istype 'reg_D,invert';


Equations

[ isolate, sample2, sample1, bselect ].oe = !0;
[ aselect, phi0, phi1, phi2, phi3, phi4 ].oe = !0;

[ isolate, sample2, sample1, bselect ].clk = pal_clk;
[ aselect, phi0, phi1, phi2, phi3, phi4 ].clk = pal_clk;

isolate := (a1 & a2 # a3) & a4 & a5 & !a6 # (!a0 # !a1 # !a2
        # !a3 # !a4) & a6;

sample2 := a0 & a2 & !a3 & a4 & a5 & !a6 # a1 & a2 & !a3 & a4 &
a5 & !a6 # a3 & a4 & a5 & !a6 # !a3 & !a4 & !a5 & a6 #
!a0 & !a2 & a3 & !a4 & !a5 & a6 # !a1 & !a2 & a3 & !a4 & !a5 & a6;

sample1 := a1 & a2 & !a3 & a4 & a5 & !a6 # a3 & a4 & a5 & !a6 #
!a3 & !a4 & !a5 & a6 # !a1 & !a2 & a3 & !a4 & !a5 & a6;

aselect := !a2 & !a3 & !a4 & !a5 & !a6 # a6 # a5 & a4 &
(a3 # a2 & (a1 # a0));

bselect := !a6 & !a5 & (a4 # a3 # a2 & (a1 # a0)) # a5 & (!a4
# !a3 & !a2);

phi0 := !a2 & !a3 & !a4 & !a5 & !a6 # !a2 & !a3 & a4 & a5 & !a6;

phi1 := !a1 & !a2;

phi2 := a1 & !a2;

phi3 := !a1 & a2;

phi4 := a1 & a2;

End pal1
```

```
// pal2.abl
//
// SCCS id: 1.1
//
// Steven J. Decker
// decker@mtl.mit.edu
//
// 5/6/1997

Module pal2

Title
'PAL2 equations
Steven J. Decker
decker@mtl.mit.edu
pal2.abl
SCCS id: 1.1
5/6/97'

Declarations

pal2 device 'P22V10C';

pal_clk pin 2;
a0 pin 3;
a1 pin 4;
a2 pin 5;
a3 pin 6;
a4 pin 7;
a5 pin 9;
a6 pin 10;
a7 pin 11;

!row_clk pin 17 istype 'reg_D,invert';
sor pin 18 istype 'reg_D,buffer';
eof pin 19 istype 'reg_D,buffer';
swsel3 pin 20 istype 'reg_D,buffer';
swsel2 pin 21 istype 'reg_D,buffer';
swsel1 pin 23 istype 'reg_D,buffer';
clkb pin 24 istype 'reg_D,buffer';
clkr pin 25 istype 'reg_D,buffer';
clka pin 26 istype 'reg_D,buffer';
```

```
clkbb pin 27 istype 'reg_D,buffer';


Equations

[ row_clk, sor, swsel3, swsel2 ].oe = !0;
[ swsel1, clkb, clkr, clka, clkbb ].oe = !0;
eof.oe = 0;

[ row_clk, sor, eof, swsel3, swsel2 ].clk = pal_clk;
[ swsel1, clkb, clkr, clka, clkbb ].clk = pal_clk;

row_clk := a0 & a1 & a2 & a3 & a4 & !a5 & a6;

sor := a0 & a1 & !a2 & !a3 & !a4 & !a5 & !a6;

eof := 0;

swsel3 := !a2;
swsel2 := a0;
swsel1 := a1;

clkb := !a0 & a1 & !a2 & a3 & !a4 & !a5 & a6;

clkr := a0 & a1 & !a2 & !a3 & a4 & a5 & !a6;

clka := a0 # a1 # !a2;

clkbb := a0 # a1 # a2;

End pal2
```

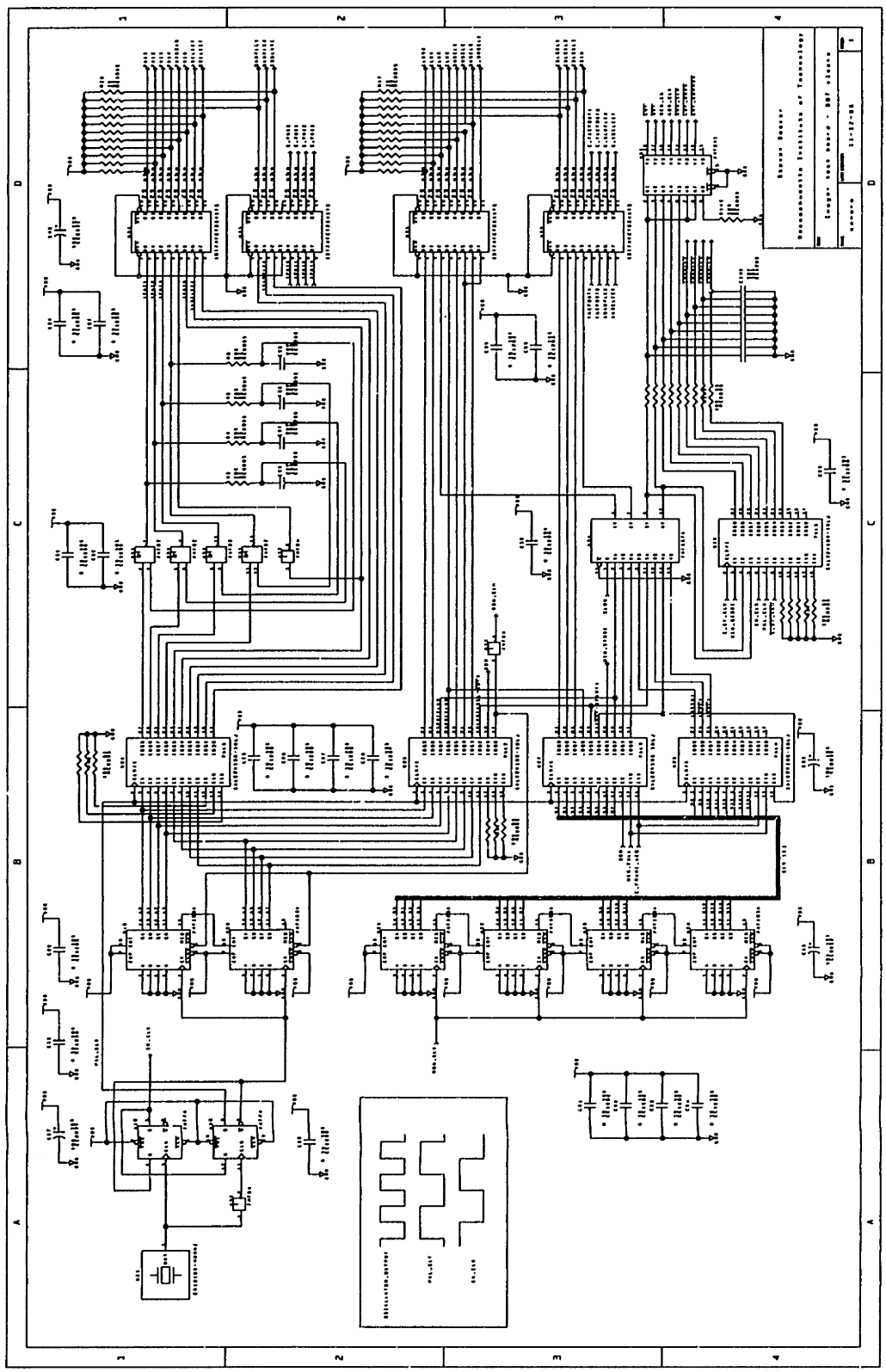
```
// pal3.abl
//
// SCCS id: 1.1
//
// Steven J. Decker
// decker@mtl.mit.edu
//
// 5/6/1997

Module pal3

Title
'PAL3 equations
Steven J. Decker
decker@mtl.mit.edu
pal3.abl
SCCS id: 1.1
5/6/97'

Declarations

pal3 device 'P22V10C';

pal_clk pin 2;
s0 pin 3;
s1 pin 4;
s2 pin 5;
s3 pin 6;
s4 pin 7;
s5 pin 9;
s6 pin 10;
s7 pin 11;
sor pin 12;
mem_full pin 13;
c_frame_acq pin 16;

clkr            pin 18;
preclkr pin 19;
preclkb pin 24;

rsin pin 17 istype 'reg_D,buffer';
mem_store pin 20 istype 'reg_D,buffer';
```

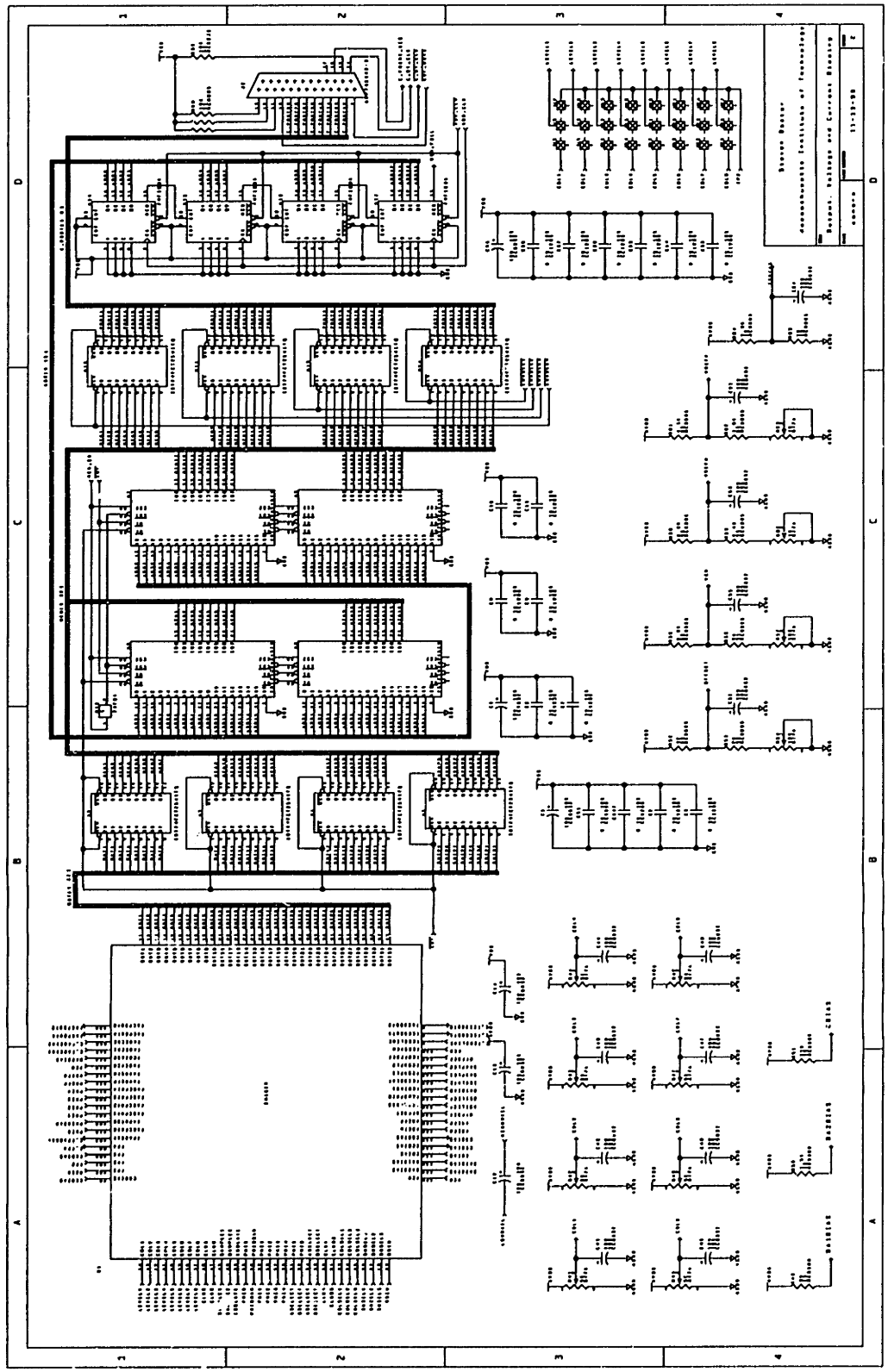
```
!sof pin 21 istype 'reg_D,invert';
!eof pin 23 istype 'reg_D,invert';
colsel3 pin 25 istype 'reg_D,buffer';
colsel2 pin 26 istype 'reg_D,buffer';
colsel1 pin 27 istype 'reg_D,buffer';

Equations

[ rsin, mem_store, sof, eof ].oe = !0;
[ colsel3, colsel2, colsel1 ].oe = !0;
[ clkr, preclkr, preclkb ].oe = 0;

[ rsin, mem_store, sof, eof ].clk = pal_clk;
[ colsel3, colsel2, colsel1 ].clk = pal_clk;

rsin := !s0 & s1 & s2 & s3 & s4 & s5 & s6 & s7;

mem_store := 1;

sof := !mem_full & !eof & c_frame_acq & (sof #
sor & s0 & s1 & s2 & s3 & s4 & s5 & s6 & s7);

eof := (sof & mem_full # eof) & c_frame_acq;

// linear imaging
colsel1 := !s7 # !s6 # !s5 # !s4 # !s3 # !s2 # !s1;
colsel2 := 0;
colsel3 := 1;

// wide dynamic range imaging (quasi-log)
// clocking order: (colsel1,colsel2,colsel3) = (0,0,0), (1,0,0),
// (1,1,0), (0,1,0), (0,1,1), (1,1,1), (1,0,1), (0,0,1).
// (1,0,1) is lowest potential, (0,0,1) is highest (reset)
// colsel1 := s7 & (!s6 # s6 & !s5 # s6 & s5 & s4 & s3 & (!s2 #
// s2 & !s1));

// colsel2 := s7 & s6 & (!s5 # s5 & !s4 # s5 & s4 & !s3 # s5 & s4
// & s3 & !s2);

// colsel3 := s7 & s6 & s5 & s4;

// wide dynamic range imaging (three-step)
```

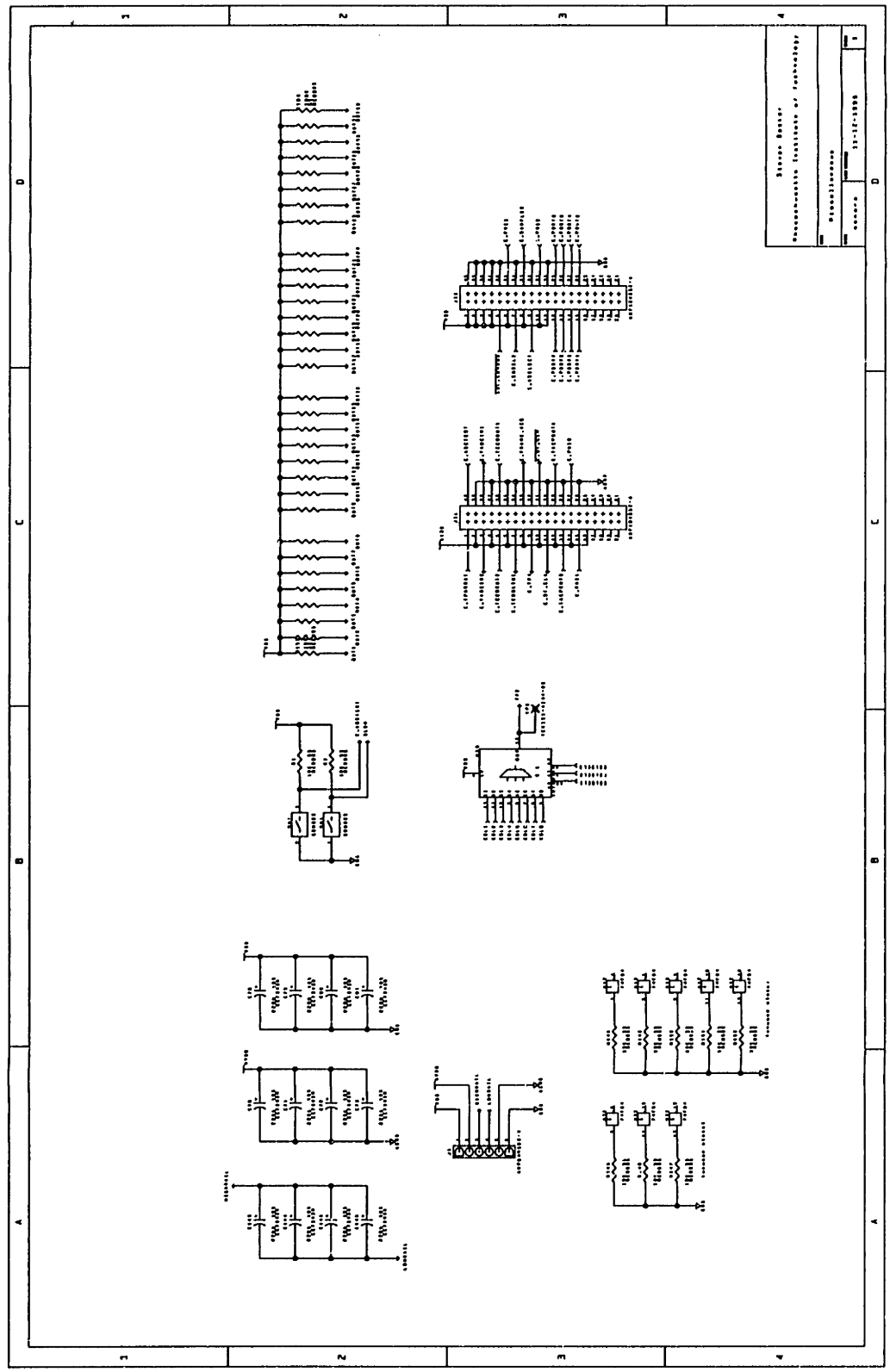
```
// clocking order: (colsel1,colsel2,colsel3) =
// (0,0,0), (1,0,0), (1,0,1), (0,0,1)
// (1,0,1) is lowest potential, (0,0,1) is highest (reset)
// colsel1 := s7 & s6 & s5 & s4 & s3 & s2 & !s1;

// colsel2 := 0;

// colsel3 := s7 & s6 & s5 & s4 & s3 & s2 & (s1 # s0);

End pal3
```

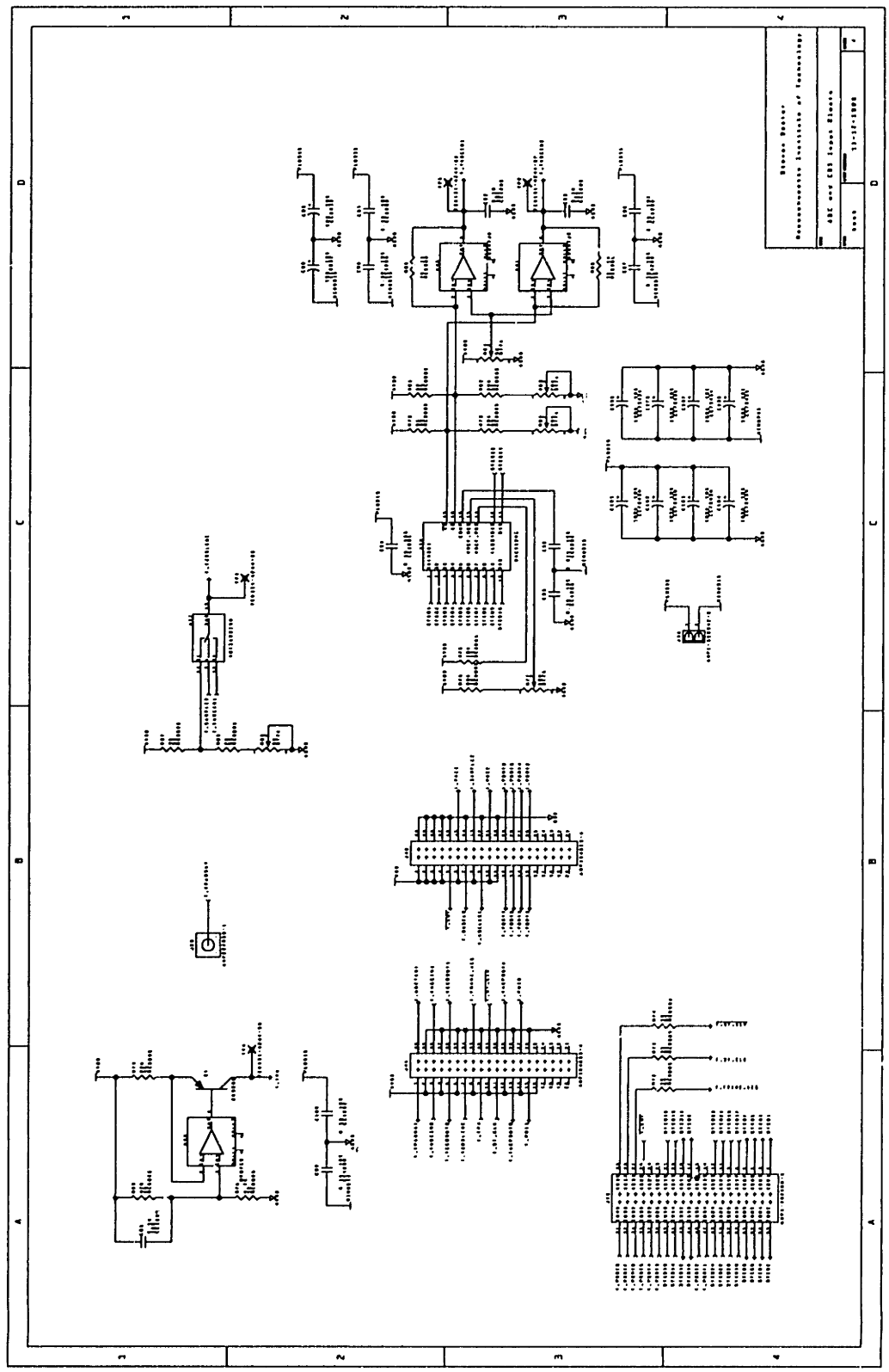
```
// pal5.abl
//
// SCCS id: 1.1
//
// Steven J. Decker
// decker@mtl.mit.edu
//
// 3/31/1997

Module pal5

Title
'PAL5 equations
Steven J. Decker
decker@mtl.mit.edu
pal5.abl
SCCS id: 1.1
3/31/97'

// Notes:
//
// 1. It is necessary to initialize the address by setting c_st_clr
//    low and clocking c_st_clk.
//
// 2. The first real data in memory is at 0x01.
//
// 3. Strawberry Tree should be toggling st_clk before board power is
//    turned on.  Otherwise, enabled drivers could fight.


Declarations

pal5 device 'P22V10C';

c_st_clk pin 2;
!sof pin 4;
!eof pin 5;
en_clk pin 6;
pal_clk pin 7;
!c_st_clr pin 9;
en1 pin 10;
en2 pin 11;
```

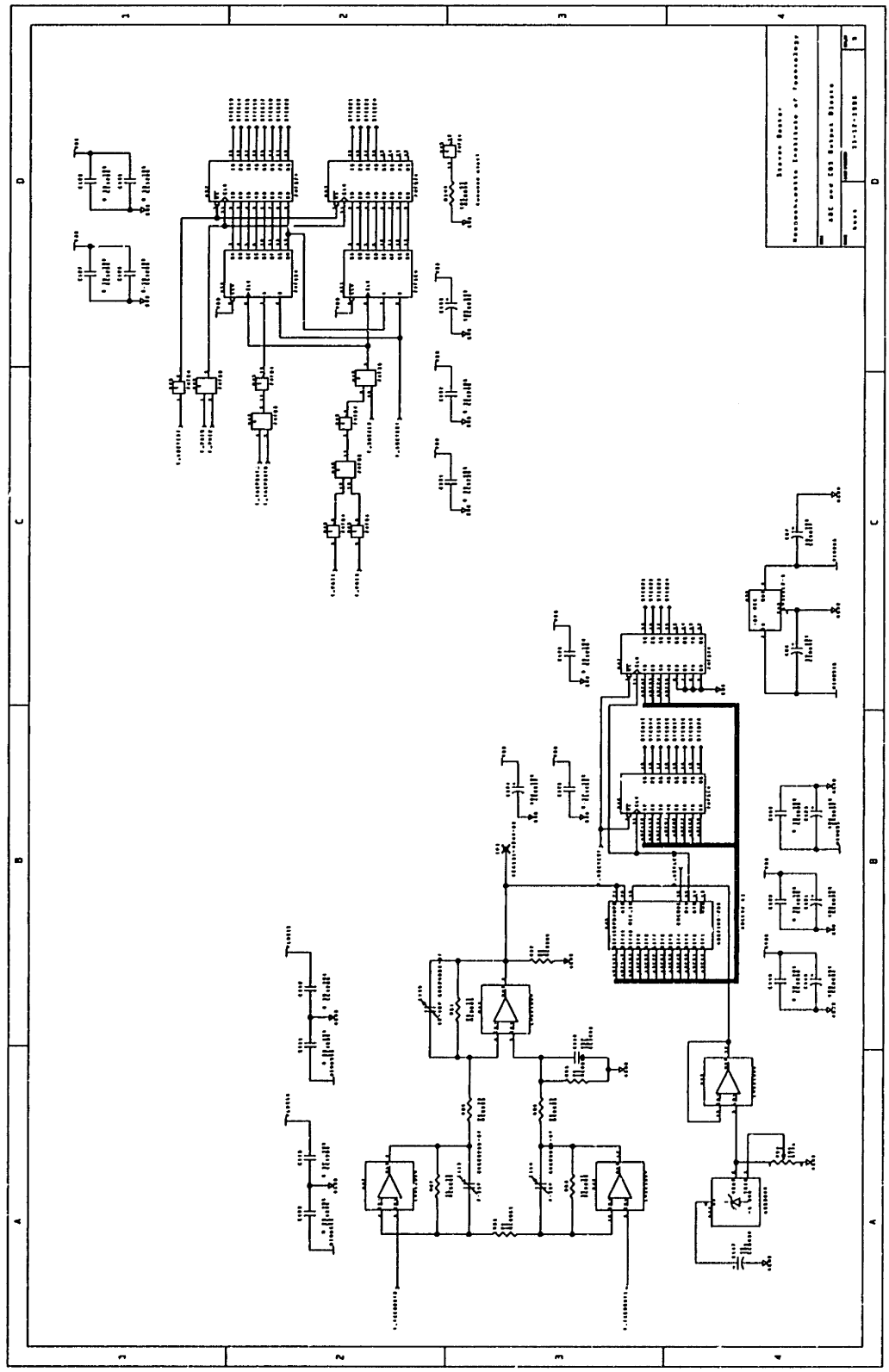
```
!enable4 pin 20          istype 'reg_D,invert';
!enable3 pin 21          istype 'reg_D,invert';
!enable2 pin 23          istype 'reg_D,invert';
!enable1 pin 24          istype 'reg_D,invert';

!adr_clr pin 25 istype 'com,neg';
adr_clk pin 26 istype 'com,pos';
mem_en pin 27 istype 'com,pos';

Equations


[ enable1, enable2, enable3, enable4 ].OE = !0;
[ adr_clr, adr_clk, mem_en ].OE = !0;

[ enable1, enable2, enable3, enable4 ].CLK = c_st_clk;

enable4 := !en1 & !en2;
enable3 := !en1 &  en2;
enable2 :=  en1 & !en2;
enable1 :=  en1 &  en2;

adr_clr = c_st_clr;

adr_clk = c_st_clk & !sof # pal_clk & sof;

mem_en = !sof # !en_clk;

End pal5
```











Appendix B

Chip Pinout and Bonding Diagram

Table B-1: Pin list, pads 1 - 72

pad #	pin	name	value	I/O/P	pad #	pin	name	value	I/O/P
1	D3	tcdsline	(1)	I	37	N4	col8	2.5 - 5 V	I
2	C2	tccbout2	(0, 5)	O	38	P3	col7	2.5 - 5 V	I
3	B1	tccbout1	(0, 5)	O	39	Q2	col6	2.5 - 5 V	I
4	D2	NC	—	-	40	P4	col5	2.5 - 5 V	I
5	E3	NC	—	-	41	N5	col4	2.5 - 5 V	I
6	C1	NC	—	-	42	Q3	col3	2.5 - 5 V	I
7	E2	NC	—	-	43	P5	col2	2.5 - 5 V	I
8	D1	NC	—	-	44	Q4	col1	2.5 - 5 V	I
9	F3	NC	—	-	45	N6	gnda	0 V	P
10	F2	NC	—	-	46	P6	clkb	(2)	I
11	E1	out1	(0, 5)	O	47	Q5	clkr	(2)	I
12	G2	out2	(0, 5)	O	48	P7	sample2	(2)	I
13	G3	out3	(0, 5)	O	49	N7	sample1	(2)	I
14	F1	out4	(0, 5)	O	50	Q6	isolate	(2)	I
15	G1	out5	(0, 5)	O	51	Q7	phi5	(2)	I
16	H2	out6	(0, 5)	O	52	P8	phi6	(2)	I
17	H1	out7	(0, 5)	O	53	Q8	Bselect	(2)	I
18	H3	out8	(0, 5)	O	54	N8	Aselect	(2)	I
19	J3	out9	(0, 5)	O	55	N9	phi0	(2)	I
20	J1	out10	(0, 5)	O	56	Q9	phi0bar	(2)	I
21	K1	out11	(0, 5)	O	57	Q10	phi3	(2)	I
22	J2	out12	(0, 5)	O	58	P9	phi4	(2)	I
23	K2	out13	(0, 5)	O	59	P10	phi2	(2)	I
24	K3	out14	(0, 5)	O	60	N10	clkA	(2)	I
25	L1	out15	(0, 5)	O	61	Q11	clkBB	(2)	I
26	L2	out16	(0, 5)	O	62	P11	phi1	(2)	I
27	M1	gnda	0 V	P	63	Q12	colsel1	(2)	I
28	N1	gndd	0 V	P	64	Q13	colsel2	(2)	I
29	M2	vddd	5 V	P	65	P12	colsel3	(2)	I
30	L3	lowrail	0 V	P	66	N11	swsel1	(2)	I
31	N2	highrail	5 V	P	67	P13	swsel2	(2)	I
32	P1	NC	—	-	68	Q14	swsel3	(2)	I
33	M3	NC	—	-	69	N12	rsin	(2)	I
34	N3	NC	—	-	70	N13	cgbias	1-5 V	I
35	P2	NC	—	-	71	P14	NC	—	-
36	Q1	NC	—	-	72	Q15	NC	—	-

(1) switched between ~2.5 V and 0.5 - 2.5 V
(2) valid lo input < 0.7, valid hi input > 4.3

Table B-2: Pin list, pads 73 - 144

pad #	pin	name	value	I/O/P	pad #	pin	name	value	I/O/P
73	M13	NC	—	-	109	C12	taphi0b	(0,5)V	I
74	N14	NC	—	-	110	B13	tadcphi0	(0,5)V	I
75	P15	NC	—	-	111	A14	tavrefp	2.5 V	I
76	M14	NC	—	-	112	B12	tavrefm	1.5 V	I
77	L13	NC	—	-	113	C11	vdda	5 V	I
78	N15	highrail	5 V	P	114	A13	tcbias2	oa2bias	I
79	L14	lowrail	0 V	P	115	B11	tadcphi2	(0,5)V	I
80	M15	vddd	5 V	P	116	A12	tadcvcm	2 V	I
81	K13	gndd	0 V	P	117	C10	tadcphi4	(0,5)V	I
82	K14	gnda	0 V	P	118	B10	tadcin2	1.5-2.5 V	I
83	L15	out17	(0,5)V	O	119	A11	tadcin1	1.5-2.5 V	I
84	J14	out18	(0,5)V	O	120	B9	tacpoutb	(0,5)V	O
85	J13	out19	(0,5)V	O	121	C9	tacpouta	(0,5)V	O
86	K15	out20	(0,5)V	O	122	A10	vdda	5 V	P
87	J15	out21	(0,5)V	O	123	A9	gnda	0 V	P
88	H14	out22	(0,5)V	O	124	B8	gnda	0 V	P
89	H15	out23	(0,5)V	O	125	A8	vdda	5 V	P
90	H13	out24	(0,5)V	O	126	C8	vcm	2 V	I
91	G13	out25	(0,5)V	O	127	C7	oa2bias	2.8 mA	I
92	G15	out26	(0,5)V	O	128	A7	vrefm	1.5 V	I
93	F15	out27	(0,5)V	O	129	A6	vrefp	2.5 V	I
94	G14	out28	(0,5)V	O	130	B7	oa1bias	1.4 mA	I
95	F14	out29	(0,5)V	O	131	B6	offset	3 V	I
96	F13	out30	(0,5)V	O	132	C6	csbias	~0.94 V	O
97	E15	out31	(0,5)V	O	133	A5	ccbias	~1.2 V	O
98	E14	out32	(0,5)V	O	134	B5	cbias	5 mA	I
99	D15	tp4 (1,2)	1-5 V	O	135	A4	tcdsvcm	2 V	I
100	C15	tp3 (1)	1-5 V	I	136	A3	gnda	0 V	P
101	D14	tp2 (1)	2.5 - 5 V	I	137	B4	tcdsphi6	(0,5)V	I
102	E13	tp1 (1)	(0,5)V	I	138	C5	tcdsphi5	(0,5)V	I
103	C14	tparray	0-1 uA	O	139	B3	tcbias1	oa1bias	I
104	B15	vddd	5 V	P	140	A2	vdda	5 V	P
105	D13	gnda	0 V	P	141	C4	tcisolat	(0,5)V	I
106	C13	gndd	0 V	P	142	C3	tcsmpl1	(0,5)V	I
107	B14	tadcphi1	(0,5)V	I	143	B2	tcoffset	3 V	I
108	A15	tadcphi3	(0,5)V	I	144	A1	tcsmpl2	(0,5)V	I

(1) Test pixel pads are row select (tp1), barrier gate (tp2), charge spill gate (tp3), and output (tp4).
(2) Output is 61.3/1.0 PMOS source follower, and should be current source biased.

Table B-3: Output bit order

	col a	col b	col c	col d	col e	col f	col g	col h
b1	1	49	2	50	3	51	4	52
b2	5	53	6	54	7	55	8	56
b3	9	57	10	58	11	59	12	60
b4	13	61	14	62	15	63	16	64
b5	17	65	18	66	19	67	20	68
b6	21	69	22	70	23	71	24	72
b7	25	73	26	74	27	75	28	76
b8	29	77	30	78	31	79	32	80
b9	33	81	34	82	35	83	36	84
b10	37	85	38	86	39	87	40	88
b11	41	89	42	90	43	91	44	92
b12	45	93	46	94	47	95	48	96

Table B-4: Imager columns converted at each output

	col a	col b	col c	col d	col e	col f	col g	col h
out 1	242	244	246	248	250	252	254	256
out 2	226	228	230	232	234	236	238	240
out 3	210	212	214	216	218	220	222	224
out 4	194	196	198	200	202	204	206	208
out 5	178	180	182	184	186	188	190	192
out 6	162	164	166	168	170	172	174	176
out 7	146	148	150	152	154	156	158	160
out 8	130	132	134	136	138	140	142	144
out 9	114	116	118	120	122	124	126	128
out 10	98	100	102	104	106	108	110	112
out 11	82	84	86	88	90	92	94	96
out 12	66	68	70	72	74	76	78	80
out 13	50	52	54	56	58	60	62	64
out 14	34	36	38	40	42	44	46	48
out 15	18	20	22	24	26	28	30	32
out 16	2	4	6	8	10	12	14	16
out 17	1	3	5	7	9	11	13	15
out 18	17	19	21	23	25	27	29	31
out 19	33	35	37	39	41	43	45	47
out 20	49	51	53	55	57	59	61	63
out 21	65	67	69	71	73	75	77	79
out 22	81	83	85	87	89	91	93	95
out 23	97	99	101	103	105	107	109	111
out 24	113	115	117	119	121	123	125	127
out 25	129	131	133	135	137	139	141	143
out 26	145	147	149	151	153	155	157	159
out 27	161	163	165	167	169	171	173	175
out 28	177	179	181	183	185	187	189	191
out 29	193	195	197	199	201	203	205	207
out 30	209	211	213	215	217	219	221	223
out 31	225	227	229	231	233	235	237	239
out 32	241	243	245	247	249	251	253	255

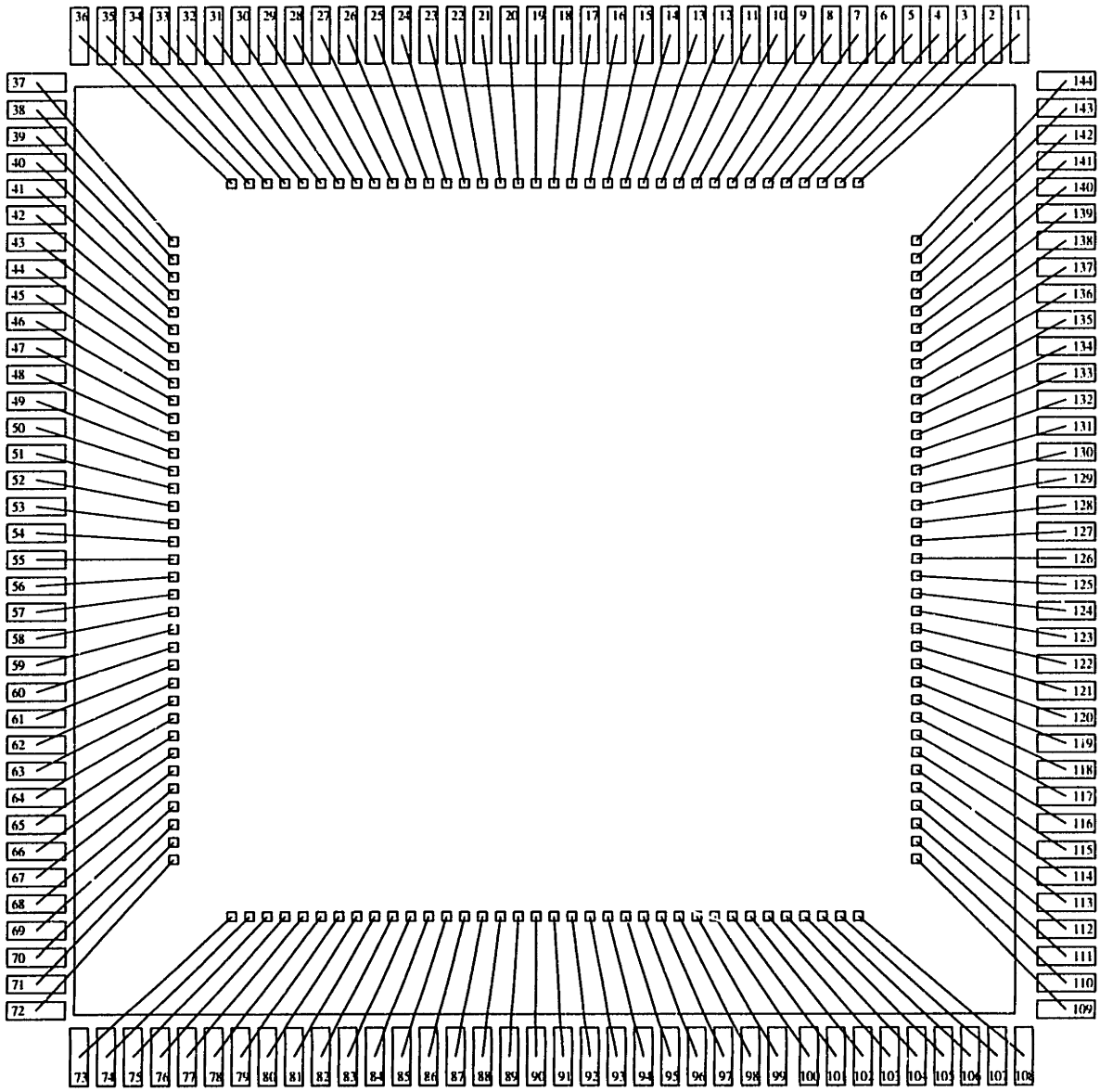


Figure B-1: Bonding diagram.

Bibliography

[1] B. Horn, H.-S. Lee, T. Poggio, C. Sodini, and J. Wyatt, "The first year of the MIT vision chip project," Tech. Rep. VLSI Memo No. 89-555, Massachusetts Institute of Technology, September 1989.

[2] D. L. Standley, "An object position and orientation IC with embedded imager," *IEEE Journal of Solid-State Circuits*, vol. 26, no. 12, pp. 1853 – 1859, February 1991.

[3] C. L. Keast, *An Integrated Image Acquisition, Smoothing and Segmentation Focal Plane Processor.* PhD thesis, Massachusetts Institute of Technology, February 1992.

[4] C. B. Umminger, *Integrated Analog Focal Plane Processing for Automatic Alignment.* PhD thesis, Massachusetts Institute of Technology, February 1995.

[5] I. S. McQuirk, *An Analog VLSI Chip for Estimating the Focus of Expansion.* PhD thesis, Massachusetts Institute of Technology, May 1996.

[6] S. Kawahito *et al.*, "A compressed digital output cmos image sensor with analog 2-d dct processors and adc/quantizer," in *1996 IEEE international Solid-State Circuits Conference*, (San Francisco, CA), pp. 184 – 185, February 1997.

[7] J. M. Hakkarainen, "A 40×40 CCD/CMOS absolute-value-of-difference processor for use in a stereo vision system," *IEEE Journal of Solid-State Circuits*, vol. 28, no. 7, pp. 799 – 807, July 1993.

[8] D. A. Martin, *ADAP: A Mixed-Signal Array Processor with Early Vision Applications.* PhD thesis, Massachusetts Institute of Technology, September 1996.

[9] J. C. Gealow, *An Integrated Computing Structure for Pixel-Parallel Image Processing.* PhD thesis, Massachusetts Institute of Technology, May 1997.

[10] S. A. Paul and H.-S. Lee, "A 9-b charge-to-digital converter for integrated image sensors," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 12, pp. 1931 - 1938, December 1996.

[11] B. Ackland and A. Dickinson, "Camera on a chip," in *1996 IEEE international Solid-State Circuits Conference*, (San Francisco, CA), February 1996.

[12] S. K. Mendis, B. Pain, R. H. Nixon, and E. R. Fossum, "Design of a low-light-level image sensor with on-chip sigma-delta analog-to-digital conversion," in *Proceedings of the SPIE Vol. 1900, Charge-Coupled Devices and Solid-State Optical Sensors III*, 1993.

[13] 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*, vol. 19, no. 1, pp. 41 - 48, February 1984.

[14] J. Mann, "Implementing early visual processing in analog VLSI: light adaptation," in *SPIE Vol. 1473 Visual Information Processing: From Neurons to Chips*, 1991.

[15] M. A. Mahowald, "Silicon retina with adaptive photoreceptors," in *SPIE Vol. 1473 Visual Information Processing: From Neurons to Chips*, 1991.

[16] U. Seger, H.-G. Graf, and M. E. Landgraf, "Vision assistance in scenes with extreme contrast," *IEEE Micro*, pp. 50 - 56, February 1993.

[17] N. Ricquier and B. Dierickx, "Pixel structure with logarithmic response for intelligent and flexible imager architectures," *Microelectronic Engineering*, vol. 19, pp. 631 - 634, 1992.

[18] B. Washkurak and S. Chamberlain, "A switched CCD electrode programmable photodetector," in *1995 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, (Dana Point Resort, Dana Point, CA), April 20-22 1995.

[19] Y. Endo, N. Harada, and O. Yoshida, "A photoelectric conversion-characteristic control method for interline transfer CCD imager," *IEEE Transactions on Electron Devices*, vol. ED-32, no. 8, pp. 1511 - 1513, August 1985.

[20] H. Komobuchi *et al.*, "1/4 inch NTSC format hyper-d range IL-CCD," in *1995 IEEE Workshop on Charge-Coupled Devices and Advanced Image Sensors*, (Dana Point Resort, Dana Point, CA), April 20-22 1995.

[21] R. Ginosar *et al.*, "Adaptive sensitivity cameras and sensors," *Advanced Imaging*, pp. 50 - 52, 199x.

[22] W. Boyle and G. Smith, "Charge coupled semiconductor devices," *Bell Systems Technical Journal*, no. 49, pp. 587 - 593, 1970.

[23] J. R. Janesick, T. Elliot, S. Collins, M. M. Blouke, and J. Freeman, "Scientific charge-coupled devices," *Optical Engineering*, vol. 26, no. 8, pp. 692 - 714, August 1987.

[24] H. L. Peek *et al.*, "An FT-CCD imager with true 2.4 x 2.4 μm^2 pixels in double membrane poly-si technology," in *1996 IEEE International Electron Devices Meeting*, pp. 35.3.1 - 35.3.4, 1996.

[25] J. T. Bosiers, E. Roks, H. L. Peek, A. C. Hleimann, and A. G. van der Sijde, "An S-VHS compatible 1/3" color FT-CCD imager with low dark current by surface pinning," *IEEE Transactions on Electron Devices*, vol. 42, no. 8, pp. 1449 - 1460, August 1995.

[26] E. R. Fossum, "CMOS image sensors: Electronic camera on a chip," in *IEEE International Electron Devices Meeting Technical Digest*, pp. 1–9, December 1995.

[27] D. K. Schroder, *Advanced MOS Devices.* Reading, MA: Addison-Wesley Publishing Company, 1987.

[28] A. J. P. Theuwissen, "CCD imaging," *Philips Journal of Research*, vol. 48, no. 3, pp. 147 - 158, 1994.

[29] J. T. Bosiers *et al.*, "A S-VHS compatible 1/3720(h) * 588 (v) FT-CCD with low dark current by surface pinning," in *1992 IEEE International Electron Devices Meeting*, pp. 5.1.1 - 5.1.4, 1992.

[30] C. H. Sequin and M. F. Tompsett, *Charge Transfer Devices.* New York: Academic Press, 1975.

[31] Texas Instruments, P.O. Box 655303, Dallas, TX 75265, *Data Sheet for TC244 786 × 488-Pixel CCD Image Sensor*, December 1991.

[32] J. T. Bosiers *et al.*, "CCD imaging," *Philips Journal of Research*, vol. 48, no. 3, pp. 159 - 184, 1994.

[33] D. N. Nichols *et al.*, "Single-chip color HDTV image sensor with two polysilicon levels and with wsi_x lightshield used for strapping vertical gates," in *1992 International Electron Devices Meeting*, pp. 101 - 104, 1992.

[34] H. Yamashita *et al.*, "A 2/3-inch 2M-pixel stack-CCD imager," in *1994 IEEE International Solid-State Circuits Symposium*, pp. 224 - 225, 1994. Full-length version in JSSC, August 95.

[35] J. T. Bosiers *et al.*, "A 1/4" true progressive scan 640(h)*480(v) FT-CCD for multimedia applications," in *1994 IEEE International Electron Devices Meeting*, pp. 29.3.1 - 29.3.4, 1994.

[36] S. Agwani *et al.*, "High resolution tri-linear colour TDI CCD image sensor with programmable responsivity gain," in *1995 IEEE International Electron Devices Meeting*, pp. 7.1.1 - 7.1.4, 1995.

[37] I. L. Fujimori, "A differential passive pixel image sensor," Master's thesis, Massachusetts Institute of Technology, 1997.

[38] R. D. McGrath, V. S. Clark, P. K. Duane, L. G. McIlrath, and W. D. Washkurak, "Current-mediated, current reset 768 × 512 active pixel sensor array," in *Technical Digest, 1997 IEEE Solid-State Circuits Conference*, pp. 182 - 183, 1997.

[39] P. K. Tedrow and C. G. Sodini, "Twin well CMOS process," tech. rep., Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, June 1989.

[40] C. L. Keast, "A CCD/CMOS process for integrated image acquisition and early vision signal processing," Master's thesis, Massachusetts Institute of Technology, September 1989.

[41] S. Mendis *et al.*, "CMOS active pixel image sensor," *IEEE Transactions on Electron Devices*, vol. 41, no. 3, pp. 50 - 52, March 1994.

[42] A. Dickenson *et al.*, "Standard CMOS active pixel image sensors for multimedia applications," in *Advances in VLSI 1995*, pp. 214 - 224, 1995.

[43] E. G. Stevens, "Photoresponse nonlinearity of solid-state image sensors with antiblooming protection," *IEEE Transactions on Electron Devices*, vol. 38, no. 2, pp. 299 - 302, February 1991.

[44] S. G. Chamberlain, W. D. Washkurak, and B. C. Doody, "CCD image sensors for high speed inspection systems," in *SPIE Vol. 849 Automated Inspection and High Speed Vision Architechtures*, 1987.

[45] W. D. Washkurak, B. C. Doody, and S. G. Chamberlain, "A high speed, wide dynamic range linear DYNASENSOR CCD detector array for acousto-optic applications," in *SPIE Vol. 972 Infrared Technology XIV*, 1988.

[46] D. Wen, "High dynamic range charge-coupled device," October 1989. U. S. Patent No. 4,873,561.

[47] J. C. Candy and G. C. Temes, "Oversampling methods for A/D and D/A conversion," in *Oversampling Delta-Sigma Data Converters*, pp. 1 – 25, IEEE Press, 1992.

[48] M. Sayag, "Non-linear photosite response in CCD imagers," October 1991. U. S. Patent No. 5,055,667.

[49] T. F. Knight, *Design of an Integrated Optical Sensor with On-Chip Preprocessing.* PhD thesis, Massachusetts Institute of Technology, 1983.

[50] Y. P. Tsividis, *Operation and Modeling of the MOS Transistor.* New York: McGraw-Hill, 1987.

[51] O. Yadid-Pecht, B. Mansoorian, E. Fossum, and B. Pain, "Optimization of noise and responsivity in CMOS active pixel sensors for detection of ultra low light levels," in *SPIE Vol. 3019*, 1997.

[52] S.-Y. Chin and C.-Y. Wu, "A CMOS ratio-independent and gain-insensitive algorithmic analog-to-digital converter," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 8, pp. 1201 – 1207, August 1996.

[53] R. D. B. Steven L. Garverick, Larry Skrenes, "A 32-channel charge readout IC for programmable, nonlinear quantization of multichannel detector data," *IEEE Journal of Solid-State Circuits*, vol. 30, no. 5, pp. 533 – 541, May 1995.

[54] R. Gregorian and G. C. Temes, *Analog MOS Integrated Circuits for Signal Processing.* New York: John Wiley & Sons, Inc., 1986.

[55] P. Aubert, H. J. Oguey, and R. Vuilleumier, "Monolithic optical position encoder with on-chip photodiodes," *IEEE Journal of Solid-State Circuits*, vol. 23, no. 2, pp. 465 – 473, April 1988.

[56] P. Duane, 1997. personal communication.

[57] R. H. Nixon, S. E. Kemeny, B. Pain, C. O. Staller, and E. R. Fossum, "256 × 256 CMOS active pixel sensor camera-on-a-chip," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 12, pp. 2046 - 2050, December 1996.

[58] S. Mendis *et al.*, "CMOS active pixel image sensors for highly integrated imaging systems," *IEEE Journal of Solid-State Circuits*, vol. 32, no. 2, pp. 187 - 197, February 1997.

[59] C. H. Aw and B. A. Wooley, "A 128 × 128-pixel standard-CMOS image sensor with electronic shutter," *IEEE Journal of Solid-State Circuits*, vol. 31, no. 12, pp. 1922 - 1930, December 1996.