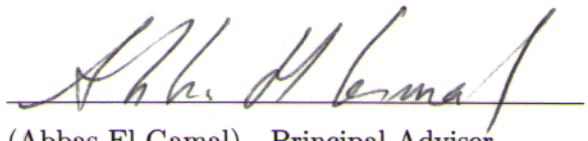
# DEVICES FOR INTEGRATED MULTI-APERTURE IMAGING

A DISSERTATION
SUBMITTED TO THE DEPARTMENT OF ELECTRICAL ENGINEERING
AND THE COMMITTEE ON GRADUATE STUDIES
OF STANFORD UNIVERSITY
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

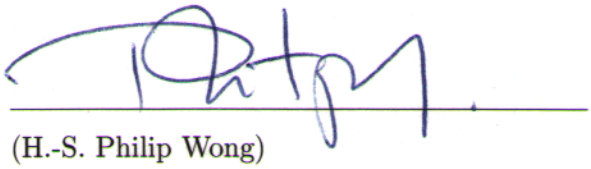
Keith G. Fife
June 2009

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



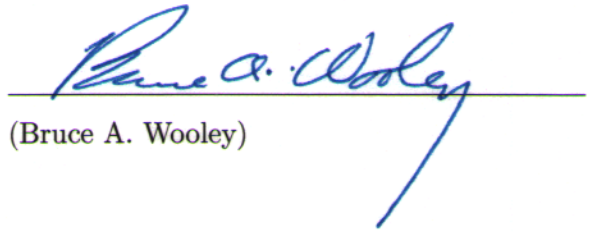
(Abbas El Gamal) Principal Adviser

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



(H.-S. Philip Wong)

I certify that I have read this dissertation and that, in my opinion, it is fully adequate in scope and quality as a dissertation for the degree of Doctor of Philosophy.



(Bruce A. Wooley)

Approved for the University Committee on Graduate Studies.

# Preface

There has been significant development of image sensors over the last decade to address issues such as sensitivity, resolution, capture rate, dynamic range, dark current, crosstalk, power consumption, manufacturability and cost. The motivation to decrease pixel size has been either to increase spatial resolution in a given format or to produce smaller formats at a given resolution for potentially lower cost. As pixel size is approaching the limits of conventional optics, improvements in resolution are diminishing. Scaling pixels beyond these limits, however, can provide new imaging capabilities beyond merely attempting to increase spatial resolution.

One consistent limitation in conventional image sensors has been that the sensing area is constrained to a regular array of photosites used to recover an intensity distribution in the focal plane of the imaging system. Although this is the most direct method of image capture, there are both practical and fundamental issues that limit the scalability or performance of these systems. This research explores an alternative approach to imaging, whereby the image sensor is partitioned into an array of apertures that form images through a distributed process.

A multi-aperture image sensor is designed with an array of apertures integrated onto a single chip. Each aperture contains its own local subarray of pixels and image forming optics. By focusing the integrated optics onto the image formed by an objective lens in a region above the multi-aperture image sensor, the apertures capture overlapping views of the scene. The correlation and redundancy between apertures, together with computation, provide several new capabilities. The most notable feature of this design, which motivates the use of submicron pixels, is that a depth map of the scene may be extracted along with the image. The accuracy in the depth calculations depends on estimating the locations of features within each sub-array of pixels. The positions of features rather than the features themselves may be estimated to a resolution higher than a diffraction or aberration limited lens can provide. Furthermore, very high resolution sensors may be made possible because

the arrays of pixels may be disjoint. This allows flexibility in readout and correction for manufacturing. Color performance is improved as neighboring pixels all contain the same filter. This design is also useful for close-proximity imaging where the objective lens can be eliminated in order to produce a flat imaging system.

Three types of submicron CCDs are implemented in single-poly 0.11$\mu$m CMOS technology to demonstrate the feasibility of multi-aperture imaging systems that produce data from distributed arrays of CCDs integrated across a monolithic substrate. Test structures comprising 16×16 pixel Frame-Transfer (FT)-CCDs with $0.5 - 0.7\mu$m pixels are fabricated under various process conditions to implement devices which operate as surface-channel, buried-channel and pinned phase buried-channel. Ripple charge transfer and single electrode charge confinement are implemented to minimize pixel pitch. Based on these designs, the first integrated multi-aperture image sensor is designed and fabricated comprising a 166×76 aperture array, each with a 16×16 buried-channel FT-CCD and per column ADCs. Measurement results suggest further scaling is possible and advantageous.

# Acknowledgement

Many people have contributed to the successful completion of this research. First, I'd like to thank the Hertz Foundation for my fellowship support during my program at Stanford. The Hertz fellowship allowed me to explore the most interesting research topics without having to commit to any external funding organization. I was able to interact with some of the brightest people through the Hertz gatherings, where we discussed many important areas of science and technology. Specifically, I'd like to express my gratitude and admiration for Dr. Lowell Wood, who not only encouraged my ideas, but also allowed me to learn from him each time we met.

I'd like to extend my gratitude to my adviser, Professor Abbas El Gamal, for encouraging me to explore this research opportunity. Throughout the project he exhibited trust in my judgment and confidence in my ability. I appreciate his guidance and collaboration on all the talks and publications we put together. He consistently encouraged me to spend my time on the most important aspects of the research and helped me to work efficiently. I also enjoyed working as a teaching assistant for him in the EE392B course on image sensors.

I also extend my gratitude to my associate adviser H. S. Philip Wong, who gave me invaluable guidance on the optimization of CCDs within a CMOS process. It was always a pleasure to meet with Professor Wong and share ideas on how to implement our work.

I also extend my gratitude to Professor Bruce Wooley for giving me the chance to work with him as a teaching assistant for EE315 VLSI Data Conversion Circuits. I've learned from his careful attention to detail and appreciated his questions during my oral examination that encouraged me to think further about some of the design choices I made.

I also thank my orals committee, Mark Horowitz, Abbas El Gamal, Philip Wong, and Bruce Wooley for taking the time out from their busy schedule to help me complete this requirement.

Without the help of TSMC, I would not have been able to turn these research concepts into physical results. From the CMOS Image Sensor Research Group, I acknowledge C. H.

Tseng, David Yen, C. Y. Ko, J. C. Liu, Ming Li, and S. G. Wuu. I am grateful for their warm reception during my visits to Taiwan and their willingness to let me work closely with them for the CMOS process modifications.

Finally, I thank my family, Katherine, Dylan, Aubrey, and Evelyn for their understanding and support throughout my graduate studies.

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

There has been significant image sensor development over the last few decades with the efforts made on charge coupled devices (CCDs) and CMOS process based devices. Several issues have been addressed such as sensitivity, resolution, capture rate, dynamic range, dark current, color filtering, cross talk, power consumption, manufacturability and cost. CMOS-process based devices have found markets in areas in which CCDs have been deficient including low cost at one extreme [1] [2] [3] and unique performance at the other [4] [5]. CMOS devices have allowed for more flexibility in overall imaging systems since digital and analog circuits are usually integrated along with the sensing elements [6]. One of the most significant constraints which remains in both CCD and CMOS image sensors is that the sensing area is made up of a regular array of photosites that is used to recover an intensity distribution that resides in the focal plane of the imaging system. The constraint is a natural result of the fact that photosites measure intensity while the optical transform is performed by a series of lenses. Typically, the lenses used in focal plane imaging systems do not approach the diffraction limit imposed by the finite aperture of the system. These lenses are usually sufficient if they achieve, at a minimum, correction for spherical aberration, chromatic aberration and distortion. As the pixel size in sensors decreases, the optical requirements become more severe. The issues involve the entire optical path including microlenses, color filters and the dielectric stack above the array. Modern CMOS processes with large dielectric stacks are not particularly well-suited for wavelength sized photosites. The incentive for reducing pixel size is often to increase resolution while decreasing system size and cost. This direction creates a set of optical issues both fundamental and practical. Traditional fixed focused systems may become impractical as the pixel size approaches the wavelength of the source. Fast diffraction-limited optics with focus control mechanisms may

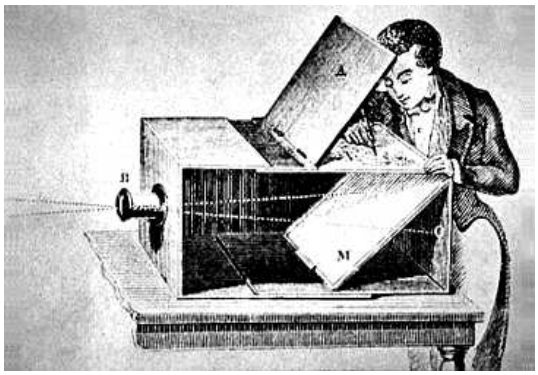
suffice for a time. The assembly and manufacturing of the optical portion of an imaging system may largely dominate the system in cost and complexity. Even if optics did not have fundamental limitations, focal plane arrays are inherently constrained by the fact that image detection is performed in the focal plane. Pixels in the array are mapped one to one with the object scene. Any defects in the array can result in defects in the resulting image. Such architectural constraints limit the scalability and performance of the image sensors used in cameras.

The purpose of this research is to explore alternative concepts to focal plane imaging by investigating devices that can be built at the semiconductor level. Breaking away from a regular array of photosites may open up applications that have not been practical in focal plane arrays. The effort is targeted at changing the requirements on the optics used to build cameras and to allow for a different approach to image detection. The idea is that the alternative methods for image recovery would reduce or eliminate some of the optical constraints that are currently driven by focal plane arrays.

## 1.1 Recent Camera History

An initial analysis of optical systems and experimentation is recorded by Alhazen in the Book of Optics (1021). Aristotle also noted that light passing through a small hole in a dark room produced an image on the wall. The *camera obscura*, Latin for "darkened room", has been used for centuries to aid artists in rendering images. Therefore, the first image capture devices required the observer to record the images by hand. In Victorian times, large public camera obscuras became popular seaside attractions. Many cameras were not made for the purpose of recording the scene at all. The first photograph was made in 1814 by Nipce by coating a pewter plate with bitumen and exposing the plate to light. Even after much refinement, these photographic cameras were simply a development of the original camera obscura. George Eastman started manufacturing paper film in 1885 and offered the first Kodak camera for sale in 1888. The camera contained a fixed-focus lens and mechanical shutter. Film technology continued to advance over the next century, producing more convenient options such as those based on silver halide. Eventually film systems would be replaced by solid-state imaging devices, which enables more recording and storage options as well as smaller form factors. A brief history of image recording technologies is given in Figure 1.1.

Solid-state imaging devices have been seriously developed since 1973 with a main

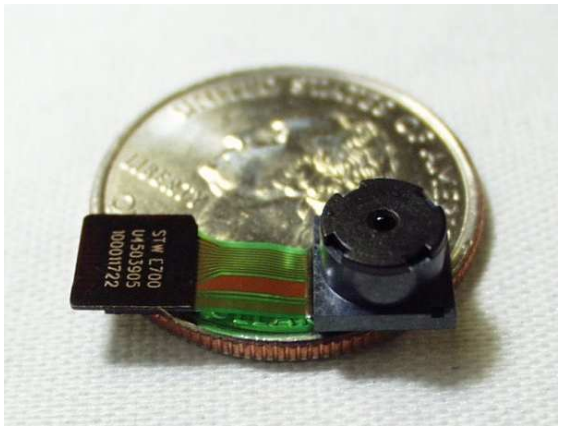


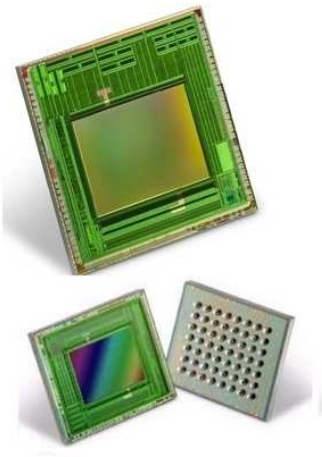
Credit: Museum of the History of Science, Oxford

(a) Camera obscura: Images recorded by hand (1015-1900).


Credit: Kodak Retina I Type 148, c. 1939. Photo by camerafiend.

(b) Film camera: Images recorded by film (1814-present).



Credit: STMicroelectronics, Switzerland

(c) Solid-state camera: Images recorded by semiconductor (1970-present).

Figure 1.1: Advances in the image recording process.

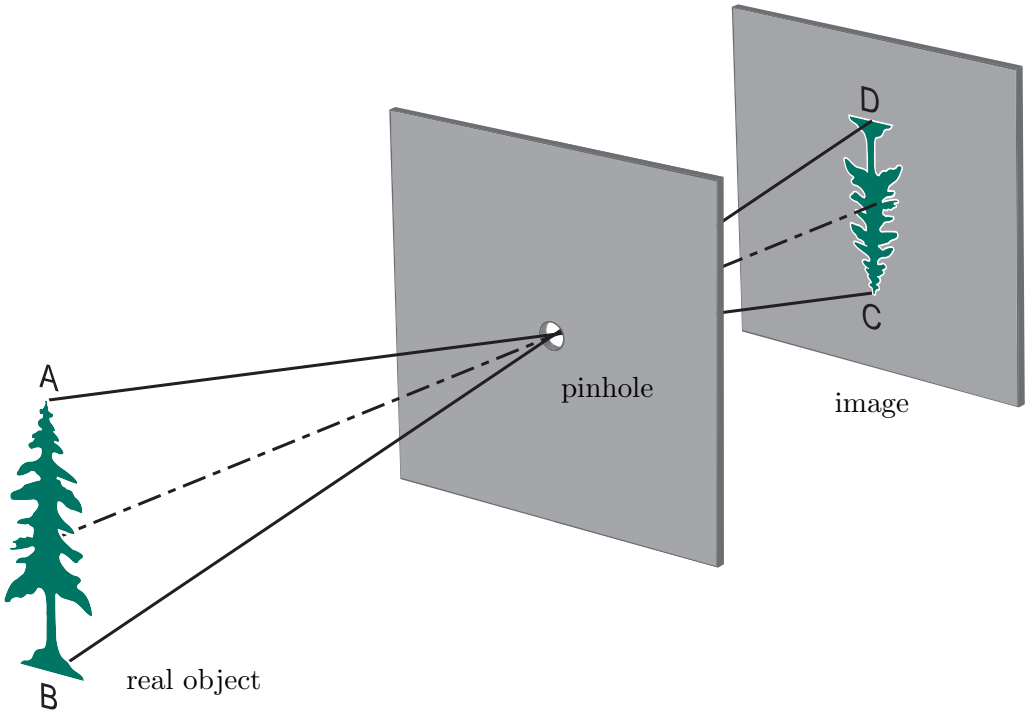
focus on CCD technology. The first devices were invented in 1969 by Willard Boyle and George E. Smith at AT&T Bell Labs. There was a long development period between the introduction of CCD devices in 1969 and the introduction of the first solid-state cameras due to the complexity of image sensor fabrication technology [7]. Between 1979 and the early 1980s, Hitachi and several others produced MOS image sensors because the device could be more easily produced in the simpler technology pushed by DRAMs. However, CCDs soon dominated Television applications, camcorders, many nonconsumer applications, and eventually digital still cameras. The 1990s would then see the revival of MOS image sensors implemented in mainstream CMOS technology. The main advantage of CMOS technology is the ability to integrate many functions on the same chip, thereby simultaneously reducing cost, power consumption, and size.

Although solid-state devices can effectively capture images, the storage technology required to record them has taken some time to develop. Many early digital camera designs made use of electromagnetic recording media to store the images. Some of these designs are described in patents by Texas Instruments (1973), Polaroid (1978) and Sony (1981) [8]. Since this time, storage technology, such as FLASH memory, has developed at such a rapid pace that cameras are no longer limited by the storage media.

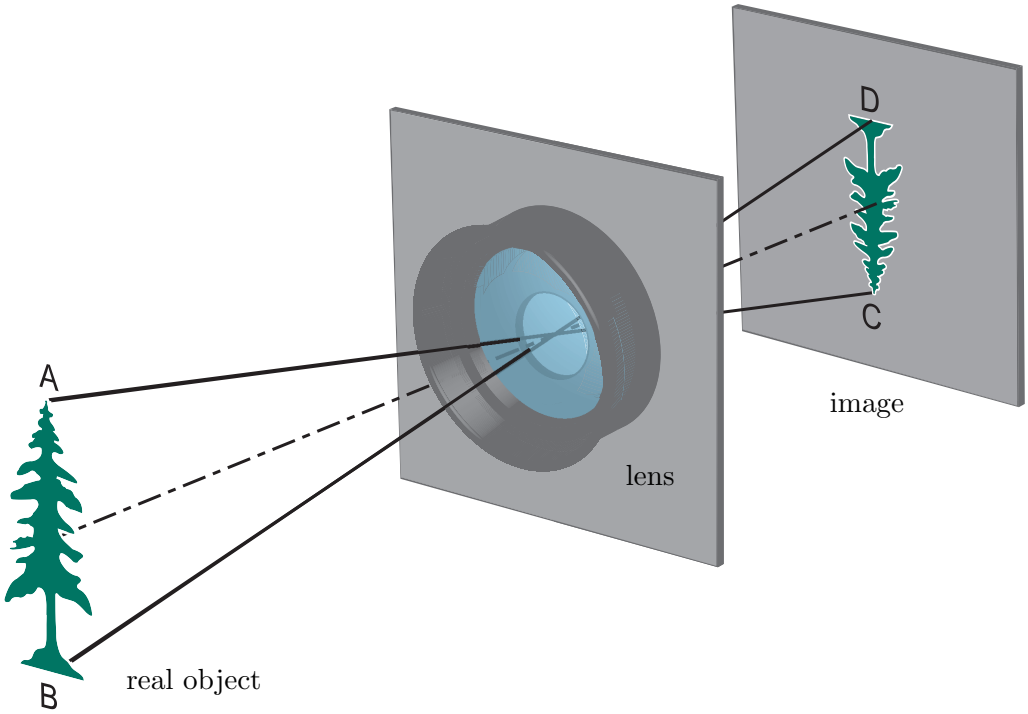
Despite the progress in the image recording processing, cameras continue to form images in the same way. Consider the configuration in Figure 1.2a. A pinhole camera can focus light by confining all rays from a scene through a single point. We can think of light traveling in straight lines such that ray tracing produces the geometric transformation from the real object to the image. Adding a lens as shown Figure 1.2b increases the intensity at any given location in the image and may produce sharper detail, but the geometric transform from point to point remains the same. Throughout time, this geometry has survived, without much deviation in the optical configuration. Although there have been significant advances in the optical elements used in cameras, these devices have mainly advanced in the way that the image is recorded.

## 1.2 Current Image Sensor Technology

There are two main types of visible image sensors available today: CCD and CMOS. Since CCD and CMOS sensors operate on similar principles, there has been some debate on which technology is better for a given application [9] [10] [11] [12]. In both types of sensors, there has been a significant effort to overcome the performance limitations of small photosites.




(a) Ray trace diagram with pinhole.




(b) Ray trace diagram with lens.

Figure 1.2: Image formation in a camera.

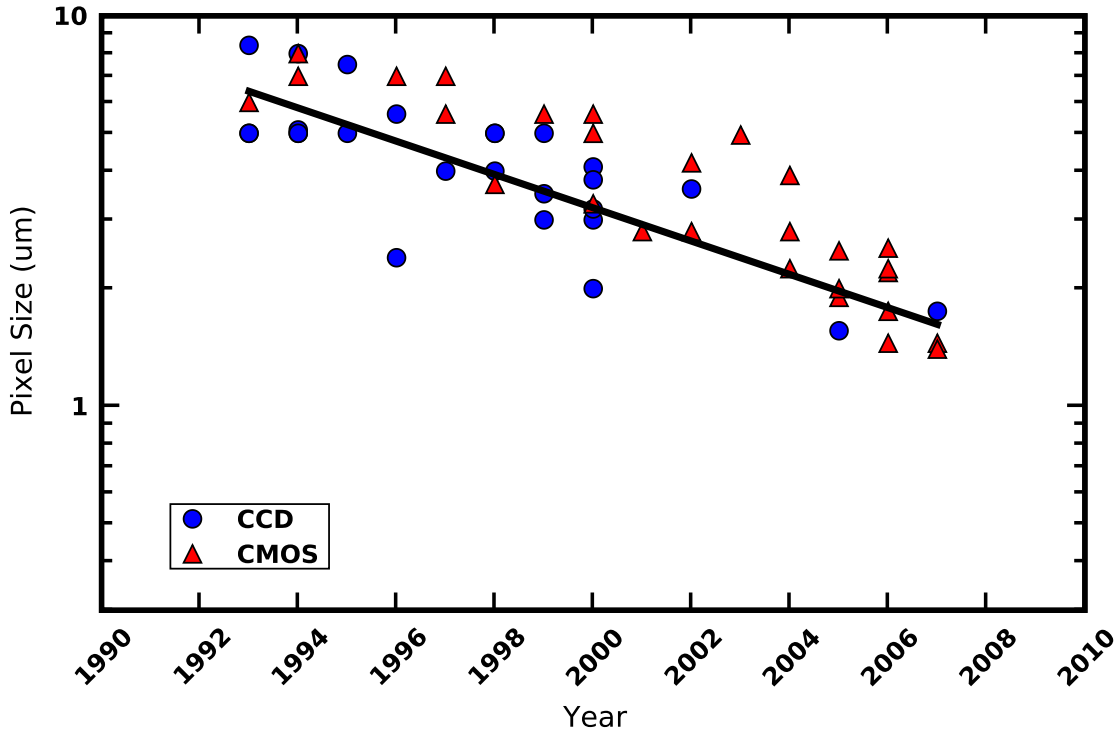



Figure 1.3: Recent pixel scaling trend.

One of the driving markets has been the camera phone business in which high resolution, low cost and small form factor are important. In general, high resolution focal plane arrays are becoming a manufacturing problem due to yield. As the pixel count increases, the number of defects per image sensor increases which results in less good die per wafer. The pixel size in papers reported over the last 15 years is shown in Figure 1.3. The pixel size tends to decrease by a factor 2 every 5 years. The benefit of pixel scaling has been either increasing spatial resolution in a given format or producing smaller formats for smaller cameras at a given resolution for potentially lower cost.

As the pixels for CMOS image sensors have become smaller, the optical and electrical crosstalk problems have increased. A cross-section of a typical CMOS image sensor with microlens array is shown in Figure 1.4 together with scenarios for on and off optical-axis pixels using objective lens with different numerical apertures (NA). Figure 1.5a shows the ray diagram for an on-axis pixel using a pin-hole objective lens. In this case, the microlens is able to focus effectively onto the photodiode area. In Figure 1.5b, the NA of the objective lens is nearly matched to the NA of the pixel, which is limited by the metal stack. The

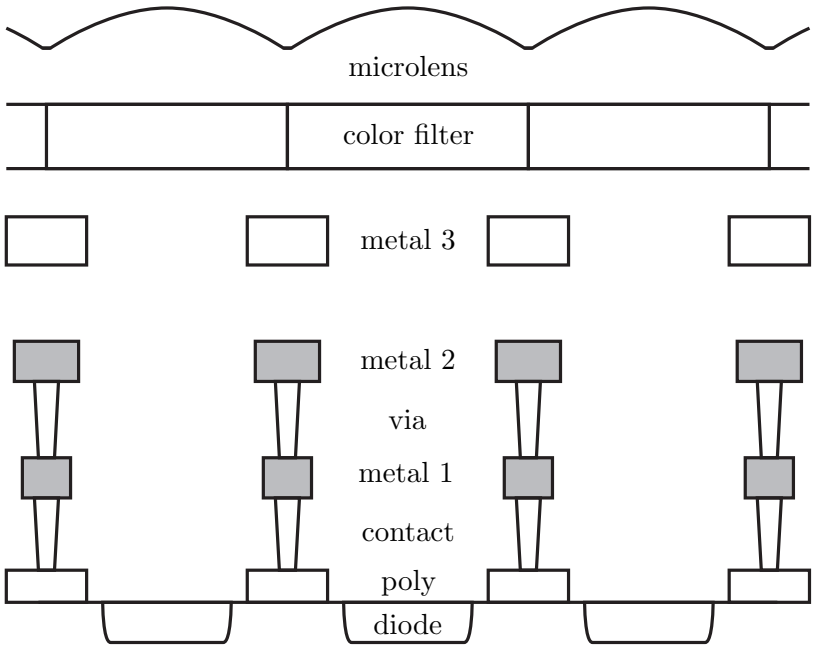



Figure 1.4: Typical CMOS pixel cross section showing the dielectric stack with color filters and microlenses.

microlens can still effectively concentrate the light onto the photodiode, albeit with a larger spot size. In Figure 1.5c, the NA of the objective lens is larger than the NA of the pixel such that the microlens is no longer effective at concentrating light onto just the photodiode area. Thus the height of the dielectric stack limits the NA of the objective lens. As pixel size is scaled, it is beneficial to increase the NA of the objective lens to gain more signal while maintaining depth of field. CMOS pixel scaling actually works against this need for a larger NA. Figure 1.5d, shows the ray diagram for an off-axis pixel using a pinhole lens. The microlens and color filter are shifted to compensate for the larger chief ray angle. In Figure 1.5e, the NA of the objective lens is matched to the NA of the pixel and significant vignetting begins to occur. This results in roll-off in signal intensity, increased crosstalk, and color gradients. In Figure 1.5f, the situation becomes even worse as rays from any particular color filter enter the wrong pixel.

A microlens study by Gennadiy et al. [13] discusses the optimum placement of the color filters and microlenses. As the pixels become smaller, the color filters and microlenses must be optimized independently for each lens that is to be used in the system. The chief ray angle and the $f/\#$ of the lens can greatly affect the optical crosstalk parameters.

There have been several approaches to compensate for the crosstalk beyond microlens

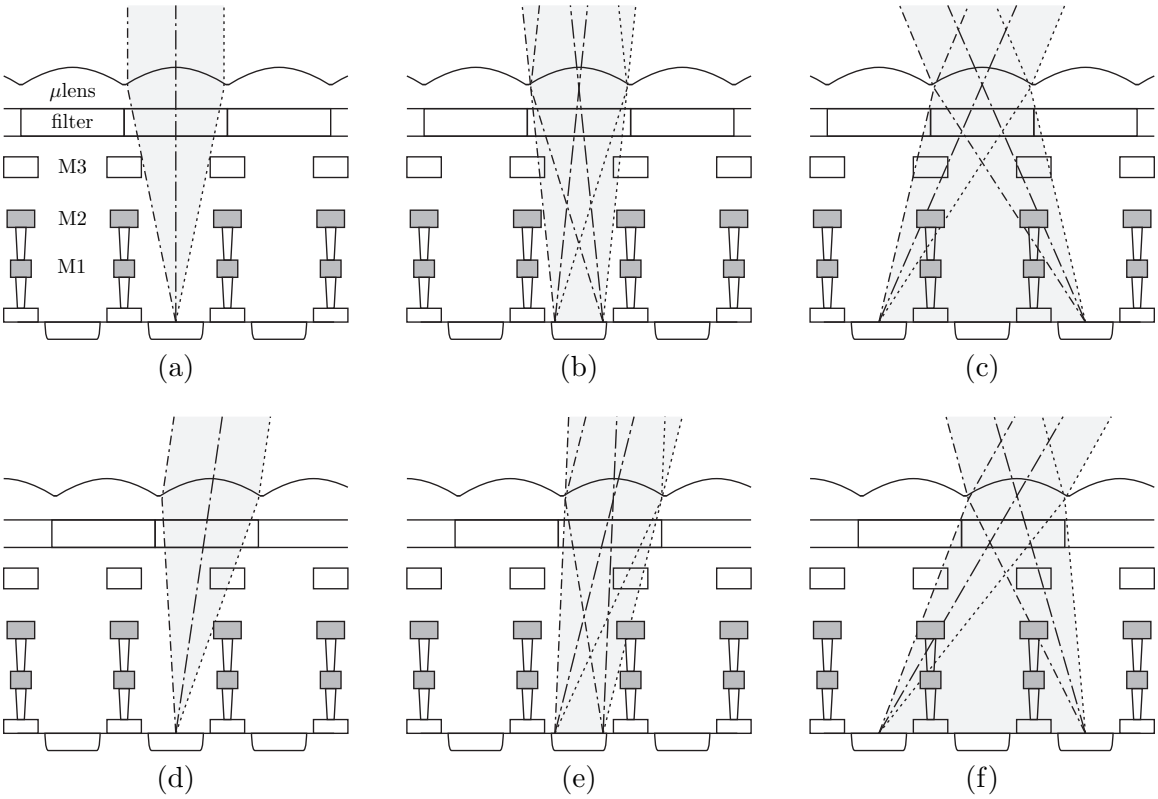



Figure 1.5: Ray diagrams for typical CMOS pixel with microlens. (a) Pinhole lens. (b) Lens NA matched to pixel NA. (c) Lens NA larger than pixel NA. (d) Pinhole lens off-axis. (e) Lens NA matched to pixel NA off-axis. (f) Lens NA larger than pixel NA off-axis.

and color filter shifting. One method reported by Ji Soo Lee et al. [14] is a measurement for the characterization of lateral crosstalk and algorithms to compensate for it based on the obtained blur model. As with many of the image processing techniques, there is a trade-off between sharpness and noise amplification. It is best to minimize the crosstalk up-front if possible.

A high-sensitivity, no-crosstalk pixel has been reported by Furumiya et al. [15]. The sensitivity is increased by using a deep p-well photodiode with an anti-reflective film consisting of $Si_3N_4$. To suppress optical crosstalk, a double metal photoshield between pixels was used. The electrical crosstalk caused by electron diffusion was suppressed with the deep p-well diode. Apparently the pixel to pixel crosstalk was suppressed to 1 percent throughout the visible range. The improvement in sensitivity from the anti-reflective film was measured to be 24 percent. Several of the Inter-Layer Dielectric (ILD) properties in CMOS sensors are discussed by Park et al. [16]. 3-D optical and electrical simulation [17] is becoming necessary in order to understand and optimize pixels for focal plane imaging.

In an effort to scale pixels to smaller sizes and take full advantage of the process design rules, TSMC has been working on an air-gap guard ring technology [18]. The idea is to simply extend the photo sensitive area of the pixel as close to the surface of the focal plane as possible. With air gaps between adjacent pixels, the refractive index differs enough to cause Total Internal Reflection (TIR) such that optical crosstalk is minimized and the usual constraints of small pixels with large stack heights can be lifted. This technology may be a good direction when considering that high levels of integration are desirable for imaging devices with high resolution and small format size. Typically, the number of metal layers should be minimized to keep the dielectric stack as thin as possible. If light guides are created with air gaps, then more metal layers may be used to enable a higher density of logic and memory without severely affecting the optical performance. Furthermore, the air gaps could help significantly reduce the off axis imaging crosstalk.

Another problem is the crosstalk that occurs within the semiconductor. Since the electron/hole pair is generated as deep as a few microns below the surface of the silicon, crosstalk due to diffusion occurs. If the substrate is uniformly doped, then crosstalk can be measured several pixels away. A model of the charge collection in the substrate is given by Chi-Shao Lin et al [19].

Scaling the pixel size also has limitations due to the amount of photons captured by each site [20]. In order to increase the amount of light available at each pixel, a low $f/\#$ lens can be used. However, the low $f/\#$ has problems due to cost, aberrations and depth

of field. All of these are in strong conflict with the reasons for using small pixels to begin with.

One of the advantages of CMOS image sensors is the use of high volume, low cost fabrication technology. When considering the future of CMOS image sensors in modern processes it becomes clear that some deviation is necessary [21]. CMOS technology is developed to optimize logic and memory for power-delay, reliability, and cost-performance. Furthermore, a large fraction of the cost of CMOS sensors is in the optical-related processes which include optical testing, optical packaging, color filters and microlenses [22]. For typical CMOS imaging systems, such as modules, a significant amount of the cost of system is in the optics and packaging. Therefore, reducing the optical complexity of imaging devices represents an important direction of research.

Another advantage of CMOS image sensor sensors is the integration of both analog and digital processing [23]. As focal plane arrays have been shifting towards CMOS from CCDs for many applications, much of the integration has been for camera on chip capabilities. It is possible to use the same concept of integration to enable new systems to perform image formation in alternative ways. This could have a performance and/or cost advantage over typical focal plane sensors.

## 1.3 Thesis Objective and Organization

As pixel size is approaching the limits of conventional optics, improvements in resolution are diminishing. Scaling pixels beyond these limits, however, can provide new imaging capabilities beyond merely attempting to increase spatial resolution. Given the current difficulties we are facing in conventional imaging systems, this thesis has the following objectives:

- Define a new imaging architecture which breaks through the current challenges.
- Develop a pixel which forms the core technology required for the new architecture.
- Develop an image sensor which implements the new architecture at the semiconductor level.

This thesis starts with background information on imaging systems and then follows through with the goals of the research.

Chapter 2 introduces computational imaging approaches that have been recently reported and are relevant to the goals of this thesis. These imaging systems attempt to modify the traditional imaging components and combine alternative optical components along with post processing to gain new benefits.

Chapter 3 describes the new imaging configuration designed for this thesis. First a description of the configuration is given and compared to other systems. Some of the applications are listed and a detailed analysis of how depth can be extracted from the system is shown. A careful look at resolution and redundancy is included to compare against conventional imaging systems. Several ideas are discussed for future work on this architecture.

Chapter 4 discusses the requirements for the pixels needed for the new architecture and details the design of three types of submicron CCDs implemented in single-poly 0.11$\mu$m CMOS technology. Test structures comprising 16×16 pixel Frame-Transfer (FT)-CCDs with $0.5-0.7\mu$m pixels are fabricated under various process conditions to implement devices which operate as surface-channel, buried-channel and pinned phase buried-channel. Ripple charge transfer and single electrode charge confinement are implemented to minimize pixel pitch.

Chapter 5 describes the design and fabrication of the multi-aperture image sensor, which comprises a 166×76 array of 16×16, 0.7$\mu$m pixel, FT-CCD subarrays with local readout circuit, per-column 10-bit ADCs, and control circuits. The image sensor is fabricated in a 0.11$\mu$m CMOS process modified for buried channel charge transfer. Global snap shot image acquisition with CDS is performed at up to 15fps with 0.15V/lux-s responsivity, 3500e- well capacity, 5e- read noise, 33e-/sec dark signal, 57 dB dynamic range, and 35 dB peak SNR.

Chapter 6 concludes the thesis with a summary of results and a discussion on the future direction of this research.

# Chapter 2

# Computational Imaging

In an effort to advance imaging systems beyond focal plane arrays that typically use independently optimized optics, an alternative direction of research has become increasingly important. Computational imaging is an approach in which the image acquisition process is shared between the optics, sensor and other hardware. A nice summary of recent work in this area is found in the Optics Express focus issue: Integrated Computational Imaging Systems [24]. The benefits of the computational imaging described there are consistent with the goals of this thesis. One paper by Mait, Athale and van der Gracht [25] outlines three main trends in computational imaging: wavefront encoding, multiplex imaging, and feature extraction. This chapter discusses both the wavefront encoding and multiplex imaging approaches after first investigating some earlier forms of optical compensation by use of signal processing.

## 2.1 Optical Compensation

One of the objectives of this thesis is to understand what can be done at the sensor level to reduce or simplify the optical requirements of the system. One basic concept is to use sub-optimal optics and introduce correction with computation. Some common forms of correction for optics include compensation for barrel distortion and cosine roll-off. Some other common techniques are found in books dedicated to image processing [26] [27]. The effectiveness of such correction, and the drive for less expensive optics, has become common enough that some image sensors now embed the functionality on chip [28].

Correction for non-ideal optical elements with the use of focal plane arrays marks early stages of computational imaging. Pushing this one step further is to design optics or the

system such that the computation produces a result or a desired property in the final image that would not normally be possible by simply correcting the optics. For example, introducing a deliberate aberration or blur to the system in which the algorithm is aware can enable a different approach to image formation. One challenge with high resolution sensors is the manufacturing yield problem due to the number of defects tolerable in an image. By blurring the image over a known spatial frequency and then recovering from the blur with computation, redundancy is enforced and pixel defects can be effectively removed and replaced with valid information. A recent work on defect-pixel recovery for high resolution image capture [29] makes use of images defocused over a 3x3 pixel space. The recovery process is performed with a Landweber-type iterative method [30]. This process attempts to recover from defective pixels and de-blur the image. With a set of known parameters in the system, it is reported that both high spatial frequency and defect recovery is possible. It is not clear how this method will work with color image recovery. The reported method claims to achieve nearly the same spatial frequency as a focused image capture. Since a color image captured with a Bayer pattern would require more blur or at least a different algorithm, it may be reasonable to expect a decrease in the final spatial frequency.

## 2.2 Wavefront Coding

A new approach to image formation called Wavefront Coding was created by Cathey and Dowski [31], developed by CDM Optics Inc, and has been presented by Kubala et al [32]. Wavefront Coding uses optics and a corresponding set of algorithms to generate and recover information about the scene. The approach involves deliberate modification of an optical image that can be post-processed in a way that ideally improves the system performance or allows for ranging information.

The incoming light rays are optically encoded and captured with a focal plane array and converted to the digital domain. The image decoding process involves the use of algorithms that are designed along with the optics. This approach can enable several different applications depending on the optical modifications. A typical problem with fast lenses is the short depth of field. A wavefront coded system can allow well-corrected systems to have extended depth of field at large apertures [33]. Furthermore, this approach can allow simple, low cost optics to be used to emulate the performance of more costly systems. A Wavefront coded system is designed for a special-purpose blur that apparently creates invariance to many optical aberrations such as spherical aberration, field curvature,

astigmatism, and chromatic aberration. Even temperature related or assembly related defocus can be corrected with signal processing. The information can also be encoded and recovered in such a way as to provide passive ranging information.

Wavefront Coded systems are different from conventional imaging systems in the use of aspheric optics that are designed with certain characteristics that mark information. Kubala et al. show how wavefront encoding reduces the optical complexity of a system by using a single aspheric lens and signal processing for recovery. A single element design is incapable of forming a high quality image without severe aberrations, which is why multiple elements are typically configured for the optical correction. A comparison is made between a two-element design, a single element design and a single element wavefront coded design. The single element wavefront coded design is able to meet the performance specifications of the two-element design by optimizing both the optics and the signal processing. The wavefront coding scheme also allows for a more compact and lower cost system.

One thing to expect from such a system is an increase in noise at a given resolution. "Noise effects are due to the need to remove the blur on the detected image. Signal processing that removes the image blur amplifies and changes the phase of the spatial frequencies. This amplification not only amplifies the spatial frequency content of the ideal image, but also the noise in the image. In practice it is this noise amplification that typically sets the limit on the amount of benefit achieved from Wavefront Coding in a particular system configuration." [32]

## 2.3 Multiplex Imaging

In multiplex imaging, optics are used to produce redundant information which is then extracted or enhanced by computation. A typical system would involve structures which contain both optical and electronic capabilities. Such optoelectronic hybridization may lead to a new methodology for compact integration of imaging systems. Coded aperture imaging, tomography, and synthetic aperture imaging are some examples of multiplex imaging.

Tanida et al. have been developing a compound imaging system they call TOMBO (Japanese for dragonfly) [34] [35] [36]. The original design objective of TOMBO was to produce a thin imaging system without the restrictions imposed by an imaging lens. Rather an array of lenslets is placed over a focal plane array. A higher resolution image is produced from several lower resolution images.

In one implementation of the TOMBO [37], a CMOS image sensor is used with 10$\mu$m

pixel pitch. The compound system consists of a microlens array, signal separator and photodector array. The signal separator and microlens array are placed over the CMOS imager chip. 10×10 unit images were captured at 8 bit resolution. Each unit consists of 50×50 pixels. The object to be captured by this system was a small transparency, backside illuminated and placed 350 mm from the microlens array. Each unit uses a small-sized high f/number microlens. The optics in this configuration are bound to the sensor. Since the microlenses are used to form images, the quality of the lenses must be higher than those typically used in focal plane arrays where they serve as concentrators to increase the effective fillfactor of each pixel.

As an extension of the TOMBO system, two methods for color imaging have been reported [38]: color separation by unit and color separation by pixel. The color separation by unit contributes to easier fabrication of color filters because the filter pitch is much wider than a single pixel pitch. In addition, there is a relaxed imaging condition because each adjacent pixel remains within the same color channel. Color separation by pixel is consistent with standard Color Filter Array (CFA) procedures for focal plane arrays. In order to compare the two methods, an object is located at twice the distance (52 cm) for the case of the color separation by unit than by case of the color separation by pixel. Those distances satisfy the condition that the units capture the color information uniformly. The longer observation length for the color separation by unit may be preferable for typical imaging systems with longer observation distances. In general a downside to the compound imaging system is that the imaging performance is strongly dependent on the object distance.

Major sources of noise on the reconstructed images from TOMBO are found to be the aberrations of the microlens and misalignment of the system components. The microlenses used in the systems so far have had minimum resolutions of approximately 3$\mu$m on the optical axis. Off axis, the value becomes tens of micrometers. These compound imaging systems require much better imaging performance than what is expected from the the conventional use of a microlens which is to concentrate the optical power. For the future research, important issues such as microlens improvement and system alignment will be considered.

Another approach to generating high resolution images from low resolution is reported by Wei and Binnie [39]. They have demonstrated an iterative algorithm for reconstruction of a high resolution image from a sequence of low resolution images captured by a CMOS sensor. The sensor is displaced in subpixel increments between successive images to generate a series of low resolution images used for the processing.

## 2.4 A Digital Lens

Since all of the data to form an image is available at the surface of the lens, it is conceivable that all of the processing could take place at that location. Fundamentally, there is no requirement for a back focal length of typical size. If the wavefront at the surface of the sensor can be completely measured and characterized, an image can be formed computationally. This would allow a range of possibilities such as extended depth of field, multiple focal points, and range detection as has been targeted by some computational imaging systems discussed above. Even without these imaging enhancements, simple image formation from a system that is not constrained by the traditional camera format is still interesting.

A lens is a very effective processor of wavefronts for image formation. To replicate this functionality with an alternative device is a challenge. The amount of data necessary to characterize a wavefront to be refocused for image formation is enormous. A system that collects all of the data first in order to form an image would be extremely inefficient compared to a typical camera. One approach is to develop a system in which the processing occurs during the detection of the wavefront. The focal point, for example, could be determined ahead of time so that the electrooptical processing that is performed is catered to those parameters and the amount of data collected is largely reduced or processed while the sensor is integrating photons.

Some computation required for a lensless imaging system using a standard focal plane array has been presented by Nicholas George in 1996 [40]. This theoretical setup uses a planar transparency with aperture illuminated by spatially incoherent light that is separated from a CCD focal plane array that records the information. The approach is to show how feasibly the image blur due to diffraction can be removed. The paper explains that the transfer function contains a modified Bessel function of the second kind which has no zeros for real values of its argument. This signifies that the image from the transparency can be recovered. With some intuition, it seems that the quality of the recovery is strongly dependent on the distance of the transparency to the detector.

# Chapter 3

# Multi-Aperture Imaging

The new imaging architecture in this work consists of an array of apertures, each with its own subarray of pixels and integrated image-forming optics as shown in Figure 3.1. The sensor readout is performed hierarchically, first at the aperture level and then globally across apertures.

The operation of a multi-aperture system is different from that of a conventional camera. Whereas in a conventional camera the objective lens focuses an image at the image sensor in the focal plane such that the pixels capture the real image, in a multi-aperture configuration, the objective lens focuses the image a certain distance *above* the sensor such that the pixel subarrays capture overlapping subimages of the real image. A high resolution 2D image can be reconstructed from the subimages via post-processing. The overlap between the subimages provides several benefits that can be achieved through more sophisticated post-processing. These include (i) obtaining a depth map of the scene, (ii) eliminating color crosstalk between different color channels by using a per-aperture color filter array, (iii) relaxing the requirements on the objective lens, and (iv) increasing the tolerance to defective pixels. In addition to these benefits, the hierarchical readout architecture of the multi-aperture image sensor makes it feasible to scale the pixel count well beyond that of conventional image sensors.

Depth resolution continues to improve with pixel scaling below the limit set by the spot size of conventional optics. The reason is that the depth of a feature is measured by the displacement between the locations of its images in different apertures, and the accuracy of localization continues to improve with pixel scaling. The trade-off between achievable spatial and depth resolutions in a multi-aperture image sensor will discussed in section 3.6. Deeply scaled pixels also enable high resolution imaging at close proximity to the sensor, which can

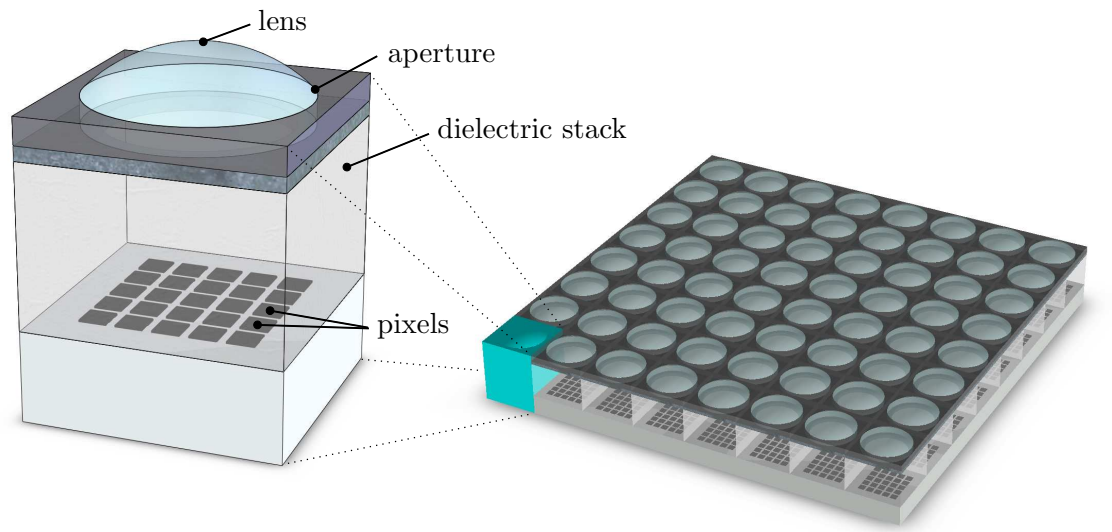



Figure 3.1: A conceptual view of a multi-aperture image sensor. Each aperture comprises a small pixel subarray and integrated imaging optics.

be useful in such applications as microscopy and in vivo imaging [41]. In this configuration, the objective optics are completely eliminated resulting in an imaging platform with very small working distance.

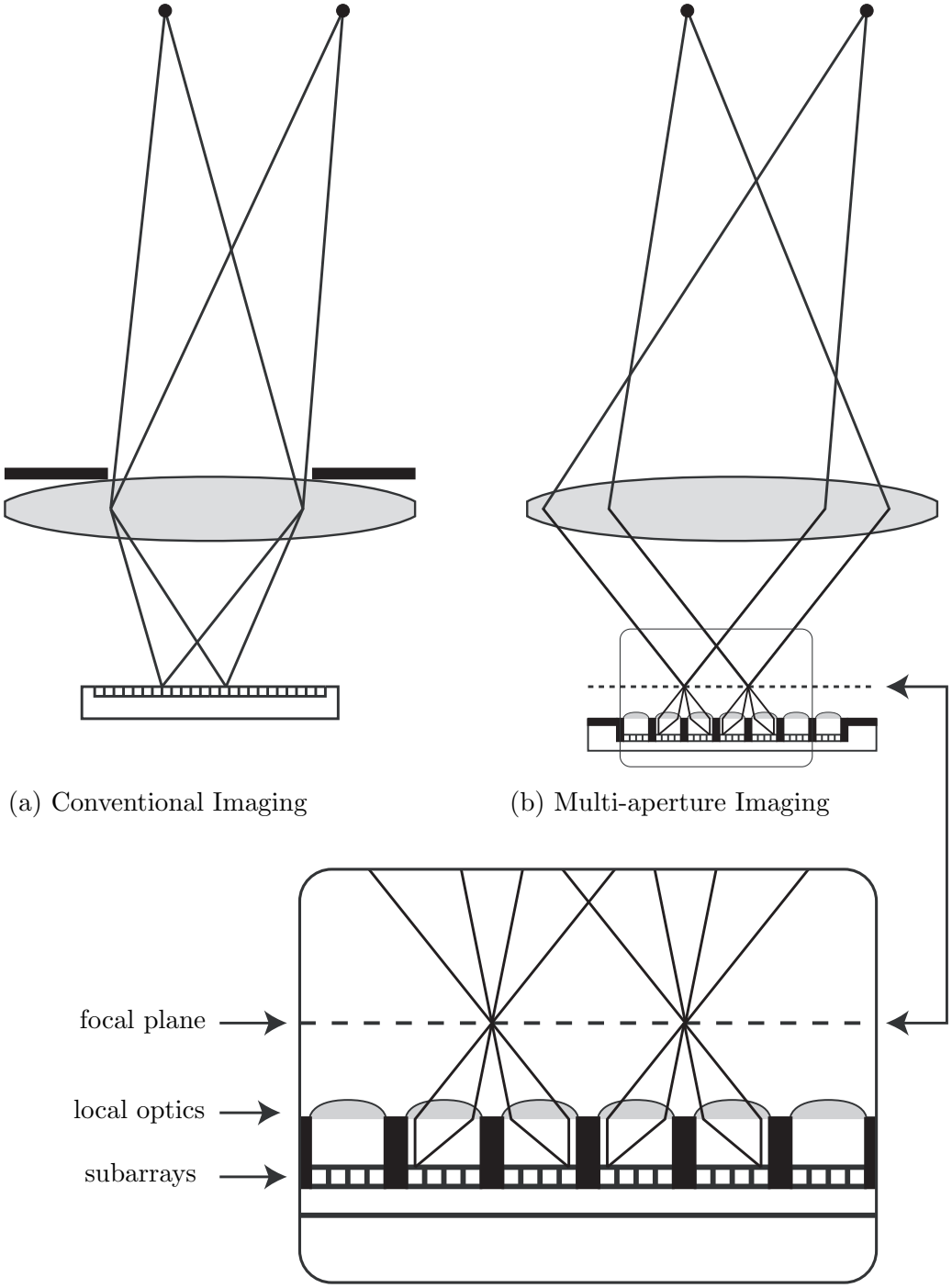
Designing scalable arrays of submicron CMOS pixels with acceptable imaging performance is challenging, however, because of the high dielectric stack height and optical occlusions resulting from the use of metal layers in the pixel. A frame-transfer charge couple device (FT-CCD) subarray is described in chapter 4 to achieve both high optical coverage and large well capacity. The image sensor is implemented in CMOS technology to enable fast multi-aperture readout and the integration of analog and digital circuits. A concern about this approach was the feasibility of implementing a CCD with acceptable charge transfer efficiency and imaging performance in a single polysilicon CMOS technology. Multiple polysilicon layers are typically used in a CCD process to create small polysilicon gap spacings, which reduce the potential barriers or pockets between electrodes [7]. As CMOS technology scales, however, the spacing between polysilicon lines becomes small enough to facilitate charge transfer between electrodes with voltage levels compatible with other CMOS circuits. Furthermore, narrow polysilicon electrodes create large fringing fields that further improve charge transfer time and efficiency. With additional implants, improvements to the standard CMOS process can be made to reduce dark current and surface traps. The feasibility of implementing a CCD in a scaled CMOS technology that is suitable for realizing

a subarray in a multi-aperture image sensor with acceptable performance is demonstrated in chapter 5.

## 3.1 Architecture

A conventional CCD or CMOS image sensor comprises a contiguous array of pixels. The objective lens focuses the image onto the image sensor (focal) plane, creating a one-to-one correspondence between points in the object space and points in the image as illustrated in Figure 3.2(a). In contrast, a multi-aperture image sensor comprises an array of pixel subarrays with gaps in between them and each pixel subarray has its own image-forming optics. The objective lens creates a focused image a certain distance above the multi-aperture image sensor plane and the local optics form secondary subimages of the image as illustrated in Figure 3.2(b). By setting the magnification of the local optics to less than one, the subimages captured by the apertures overlap. As such, each point in the object space is mapped into several points in a group of neighboring apertures. The captured subimages can be combined to form a 2D image and a depth map of the scene. Note that the objective lens in our system has no aperture from the perspective of the aperture array. This allows for a relatively complete description of the wavefront in the focal plane. The amount of depth information that can be extracted depends on the total area of the objective lens that is covered by the aperture array and ultimately on the accuracy in the localization of features.

While the multi-aperture system is similar in structure to the plenoptic system described in [42], which employs a separate microlens array on top of an image sensor, there is a key difference between the way these two systems operate. In the plenoptic system, the objective lens is focused onto the microlens array and the microlens array is focused onto the system aperture. Each microlens spreads out the incident rays to the pixels behind it, which provides information about their direction as well as intensity. While the spatial resolution of the plenoptic system is limited to the microlens count, the information about the directions of the rays can be used for a number of applications such as range finding and perspective. Note that the plenoptic system contains only one aperture (that of the objective lens), which is imaged by each of the microlenses. The useful size of each pixel in this configuration is limited by microlens aberrations and fundamentally by diffraction. In contrast, our system captures less information about the wavefront but achieves higher spatial resolution than the aperture count because each aperture captures




(a) Conventional Imaging (b) Multi-aperture Imaging

Figure 3.2: Comparison of conventional versus multi-aperture imaging.

a focused subimage with magnification greater than the reciprocal of the number of pixels in either the horizontal or vertical directions. Depth is extracted by sampling each point in the focal plane from multiple perspectives. In this configuration, pixels smaller than the spot size of the local optics are useful because they increase the localization accuracy of features and thus the displacement between the locations of the same feature in different subimages. Our image sensor architecture is also similar to that of the TOMBO compound-eye [34] whose purpose is to realize a compact, thin camera with a total resolution exceeding that of an individual aperture. The TOMBO's spatial resolution is largely dependent on object distance, while our multi-aperture image sensor confines the imaging to a tight region behind the objective lens to enable high spatial resolution in both 2D and 3D imaging.

Figure 3.3 illustrates the process of image formation and the reconstruction of a 2D image in a multi-aperture imaging system. The figure shows the virtual image formed at the focal plane and the resulting subimages captured by the apertures. Note that a feature, such as the bird's eye, appears in different locations within 16 different apertures. By appropriately shifting the subimages and combining them, a 2D image can be reconstructed. In addition to the ability to form a complete 2D image of the scene, the overlap between the subimages captured by the apertures provides other benefits. In the following sections, we discuss two of these benefits in some detail.

### 3.1.1 Depth

An example of how depth information is obtained using a multi-aperture imaging system is depicted in Figure 3.4. The sphere represents an object in close proximity to the sensor. This object could be a real object or a virtual object as it appears in the focal plane of the objective lens. A point on the sphere appears at different locations in several aperture subimages. For example, the point shown in the figure appears at the locations labeled 1, 2, and 3 in the figure. Note that if the object is close to the image sensor, the displacement between the three locations is large, while if the object is farther away, the displacement is small. Thus in general, features close to the image sensor correspond to large displacements while objects farther from the sensor correspond to small displacements. A depth map can be obtained by judiciously combining this information for each feature of interest.

From the above discussion, it is clear that to accurately estimate the depth of features within a scene, it is necessary to accurately measure the locations of each feature within the subimages. We now explain why depth resolution improves with pixel scaling beyond the spot size limit of the optics. First, note that spatial resolution is limited by the degree

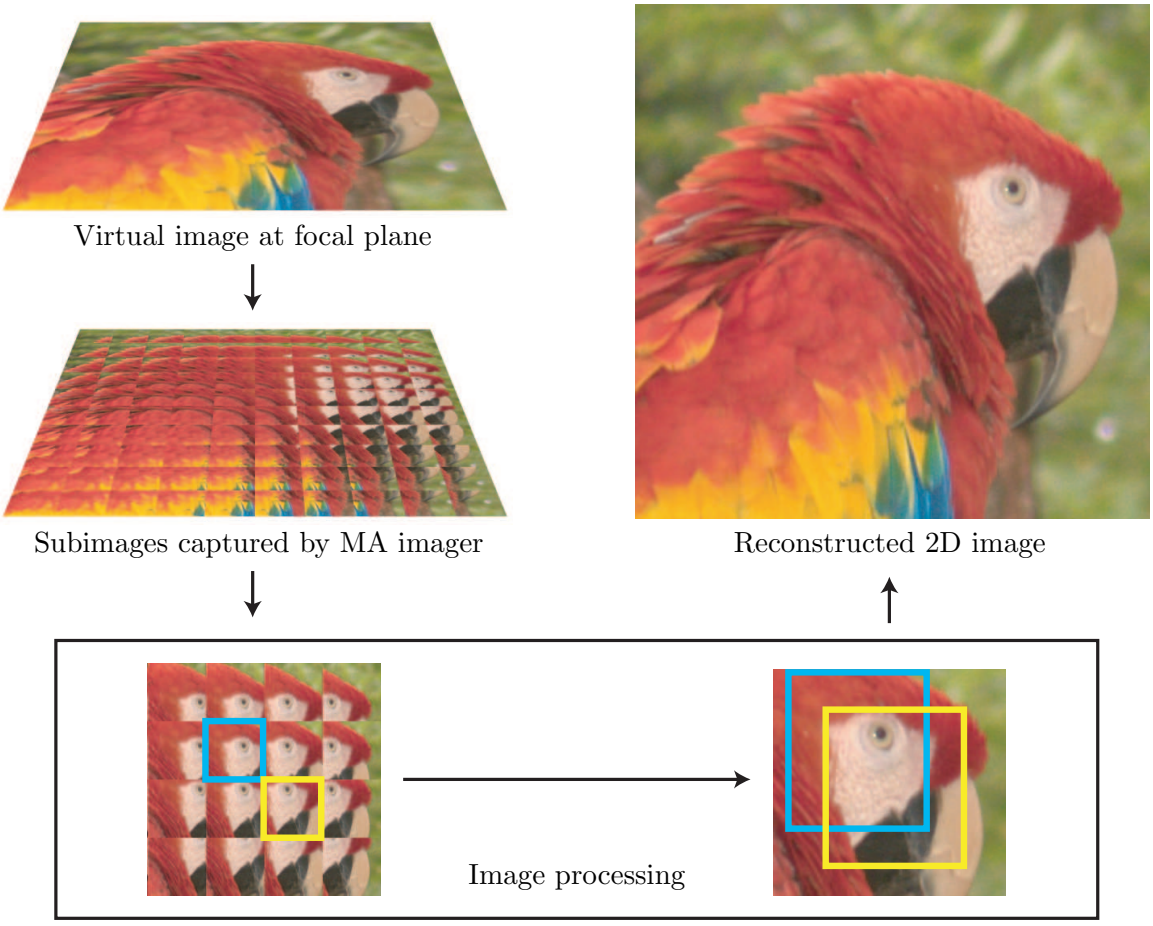
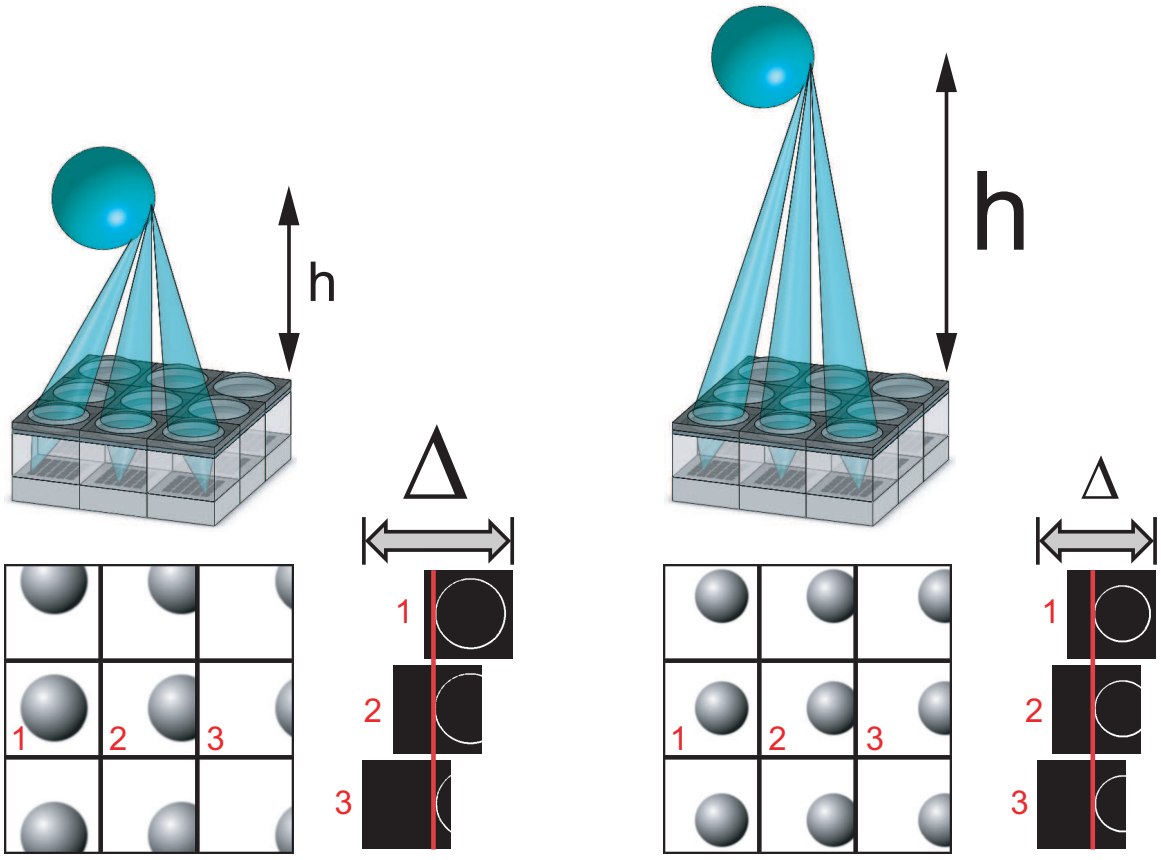



Figure 3.3: Reconstruction of a 2D image from multi-aperture subimages.




(a) (b)

Figure 3.4: Example of how depth information is obtained from displacement of feature locations in separate aperture subimages. (a) Object close to sensor corresponds to large displacement. (b) Object at greater distance from sensor corresponds to smaller displacement.

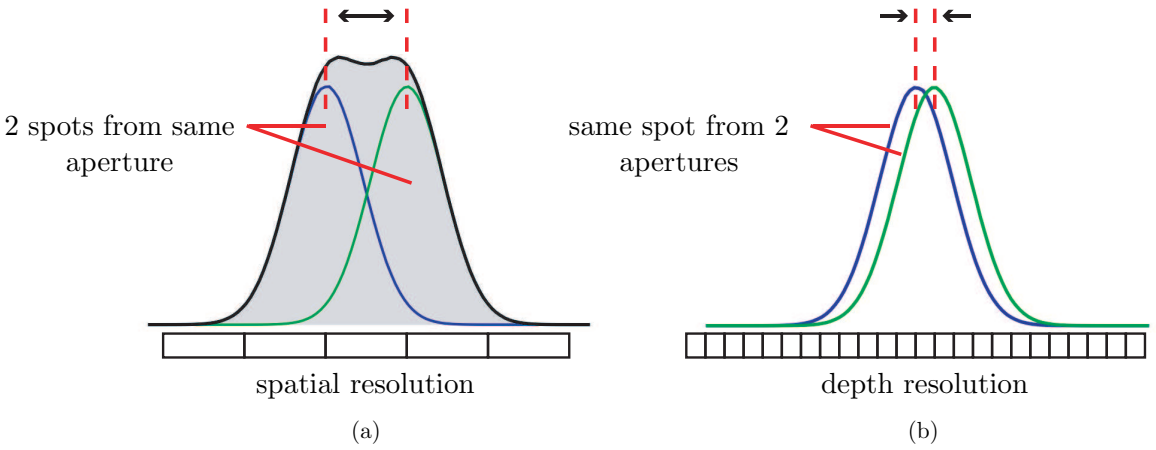



Figure 3.5: Comparison of spatial resolution and depth resolution.

to which features within the same aperture can be distinguished from each other. As the distance between features in the focal plane become smaller than the spot size as illustrated in Figure 3.5(a), they can no longer be resolved. A spot size is typically limited to a few microns in a conventional camera. On the other hand, depth resolution is determined by the displacement between the locations of the same feature in *different* subimages. These locations can be resolved to dimensions smaller than the spot size of the optics (see Figure 3.5(b)). Such precise localization was demonstrated to nanometer scale accuracy using large pixels under magnification in [43, 44]. The accuracy of localization depends on the number of photons captured for a given feature and on the accuracy of estimating the shape of the feature itself. When the pixel size is large, it is difficult to determine the shape of the feature. When the pixel size is smaller than the feature size, the shape of the feature and its location can be better determined as long as we collect enough photons. Since the localization of features is dependent on the signal to noise ratio, it is necessary to keep the read noise of the pixel low so that the read noise accumulated from several pixels sampling a single feature does not limit the location accuracy.

### 3.1.2 Color

One of the most important benefits of the multi-aperture image sensor is the ability to perform high fidelity color imaging. In a conventional image sensor, color information is obtained using a color filter array (CFA) deposited on top of the pixel array [45]. The most commonly used CFA is the RGB Bayer pattern, where a 2×2 pixel pattern

consisting of one red, one blue and two green filters arranged diagonally, is repeated over the pixel array. Such CFA configuration results in severe color crosstalk as pixel size is scaled due to scattering from the relatively thick dielectric stack and diffusion of carriers in the substrate. This crosstalk results in additional noise in the final image obtained after performing color demosaicing and correction. Furthermore, pixels far off-axis may produce color gradients for several reasons including pixel layout asymmetry. This color crosstalk problem makes further pixel size scaling without a corresponding reduction in dielectric stack height problematic. This imposes severe constraints on the efficient implementation of imaging system-on-chips.

The multi-aperture configuration offers an alternative solution to color imaging whereby pixel size can continue to scale independent of the height of the dielectric stack. This makes it possible to simultaneously scale pixel size and increase the number of metal layers to attain high logic integration density. To achieve these benefits, a per-aperture (instead of a per-pixel) color filter array such as an RGB Bayer pattern is employed. In this configuration, crosstalk between pixels is restricted to be of the same color. The magnification of the local optics is set so that each point in the scene is imaged by apertures with all three separate color filters. Some of the loss in spatial resolution due to this overlap is compensated for by imaging all three colors at each effective pixel location instead of performing color demosaicing. In our implementation, we require a local magnification of at least 1/4, which reduces the effective spatial resolution by a factor of 16. Figure 3.6(a) illustrates this new color imaging idea. The object A is focused to virtual image B. The local optics magnification is set to capture subimages at $C_R$, $C_G$, and $C_B$, each having a different color.

## 3.2 Analysis

The proposed image sensor comprises an $m \times n$ aperture array, each with its own local optics and a $k \times k$ pixel array and readout circuit (see Figure 3.7). The local optics are implemented in the dielectric stack of the integrated circuit using refractive microlenses or diffractive gratings patterned in the metal layers. The creation of independent apertures and localized pixels allows for aggressive pixel scaling, which is key to achieving high depth resolution. In this section, the image exaction process is described along with the analysis of depth and spatial resolution for the proposed architecture.

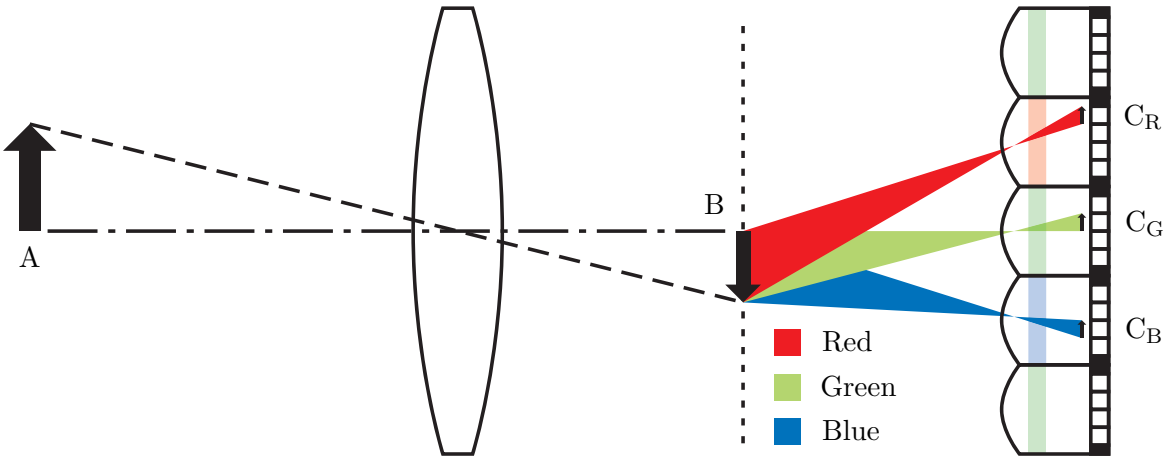



Figure 3.6: Chief ray diagram for color imaging with the multi-aperture configuration showing object at A, primary virtual image at B, and secondary images split into color channels at $C_R$, $C_G$, and $C_B$.

### 3.2.1 2D and 3D Image Extraction

Depth information is obtained from the disparity between apertures. Figure 3.8a depicts the chief ray traces for an object as it is imaged from the apertures behind the objective lens. The circle below the diagram shows the location at which the chief ray pierces the objective lens. As the object moves back and forth, the object in the focal plane (above the sensor) moves back and forth with some attenuation in magnitude governed by the lens law. The movement of the object in the secondary images formed by the local optics is lateral. Therefore, the amount of lateral displacement between multiple apertures corresponds to the depth of the object. With several apertures accurately placed with respect to each other, the correspondence between them becomes quite reliable. The marginal ray traces for the same point as seen from the two different apertures are shown in Figure 3.8b-3.8d. The circle below each diagram shows the area of the objective lens that is used by each aperture. As can be seen, a virtual stereo pair is projected up to the plane of the objective lens. The characteristics of the apertures remain constant across the array without spatial compensation, especially as the objective lens maintains telecentricity.

At nominal object distance, only a small number of apertures sample any given point in the object plane. As the object moves to further distances, more apertures capture its information. Therefore, both the position of objects within each aperture and the total number of apertures imaging the same point are indicators of depth. Since the redundancy between apertures is localized across the focal plane, spatial resolution continues to scale

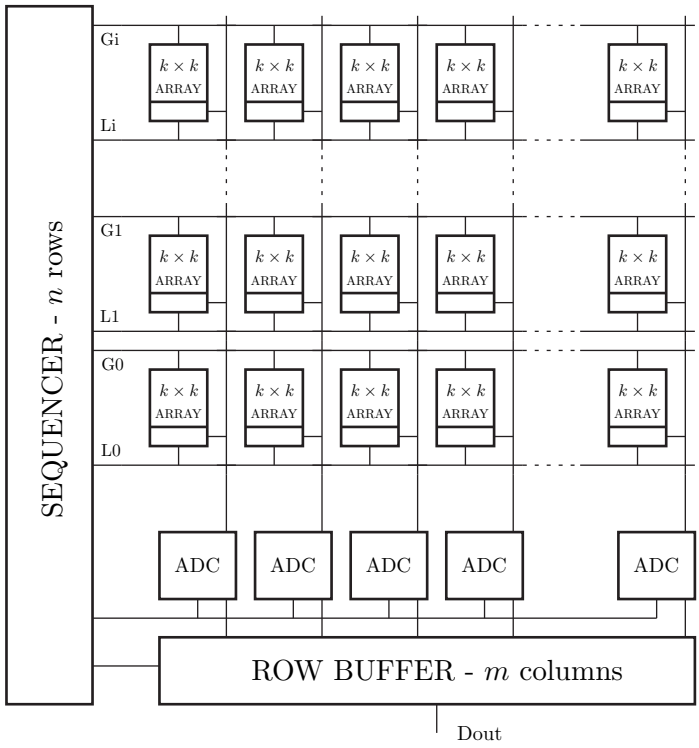


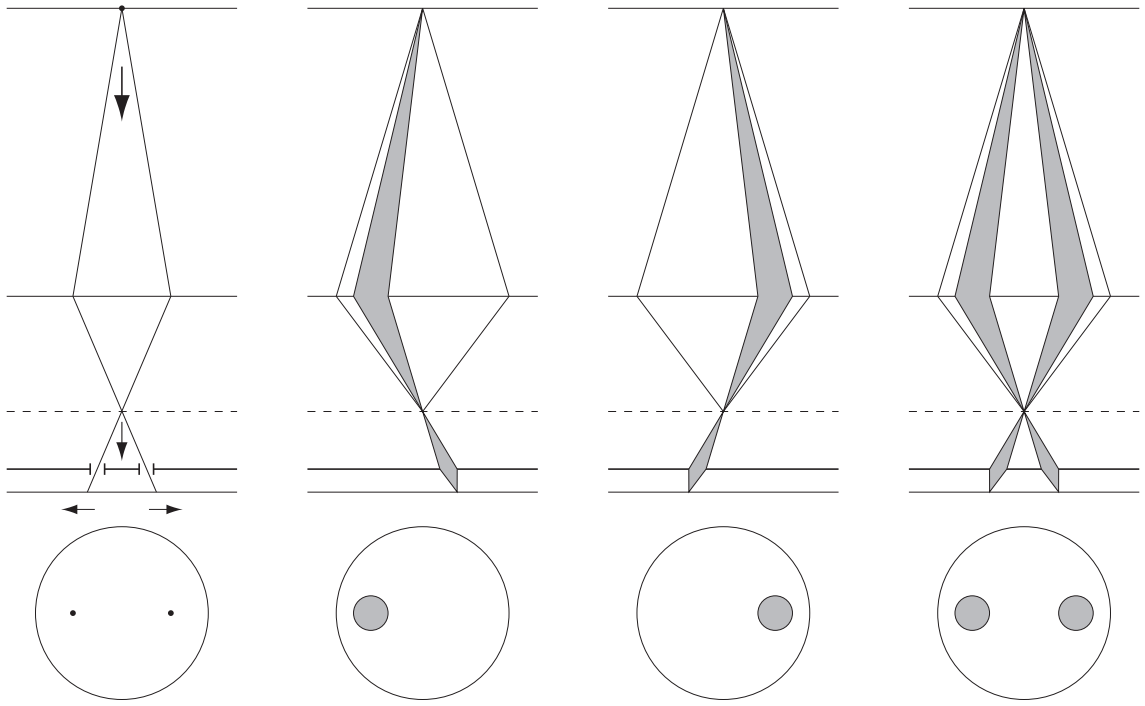

Figure 3.7: Block diagram of integrated sensor.

by adding more apertures. To increase depth resolution, as we shall see, pixel size is scaled down even below the diffraction limit. While it is difficult to scale pixels to this level in a large, uniform array, the fact that the pixels in our sensor are grouped into smaller disjoint arrays facilitates such aggressive pixel scaling.

In Subsection 3.2.2 we quantify the depth of field for our system and in Subsection 3.2.3, we establish the relationship between object distance and the displacement at the sensor surface. In Subsection 3.2.4, we discuss the dependency of the available 2D and 3D spatial resolution on pixel size and local magnification. We assume an ideal, diffraction limited optical system and ideal sensor characteristics. Of course, such nonidealities should be considered when computing the real spatial resolution limits.

### 3.2.2 Depth of Field

To evaluate the depth of field, consider the diagram in Figure 3.9. Considering the parameters defined in the figure, define the distance $E = B + C$ and the magnification factors $M = B/A$ and $N = D/C$. Because we fix the distance $E$ for a given object range,



(a) Chief rays for a pair of apertures (b) Left virtual objective aperture (c) Right virtual objective aperture (d) Virtual stereo view

Figure 3.8: Virtual aperture views.

the other variables $B$, $C$, $D$, $M$, and $N$ are all driven by the object distance $A$. Given a nominal object distance $A_0$, we denote the other parameters by $B_0$, $C_0$, $D_0$, $M_0$, and $N_0$. As $A$ varies, the distance $E$ can be adjusted to achieve the desired local magnification $N_0$ for the secondary image focused at $D_0$. This is similar to adjusting the focus in a conventional camera. The distance $D_0$ is approximately equal to the dielectric stack height of the fabrication process, or the nominal distance to the secondary focal plane from the local optics. Thus, given the stack height $D_0$, the focal length $g$ is set during fabrication to meet the desired $N_0$ value. To illustrate the results, we assume that $f = 10\text{mm}$, $A_0 = 1\text{m}$, $D_0 = 10\mu\text{m}$, and $g = 8\mu\text{m}$. These parameters yield a nominal magnification factor of $N_0 = 1/4$. This value is chosen to achieve the desired amount of overlap between aperture views as detailed in later calculations.

We solve for $D$ as a function of $A$ by fixing the parameters to meet the nominal magnification factor $N_0$. To characterize the depth of field, we find the deviation in $D$ from the nominal position $D_0$ where it is in best focus. Since the local optics collect light

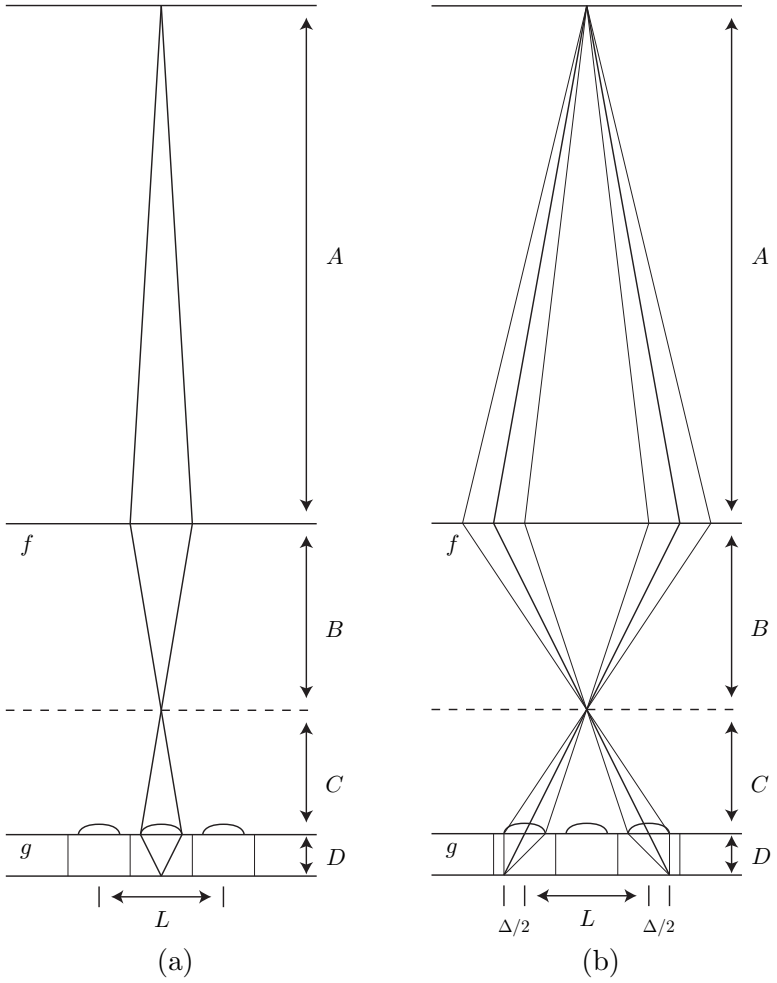



Figure 3.9: Diagrams for computing depth.

across the entire aperture, the focus is degraded with deviation in $D$. By the lens law,

$$1/f = 1/A + 1/B, \text{ and } 1/g = 1/C + 1/D. \tag{3.1}$$

Using the magnification factors $M$ and $N$, we solve for $B$ and $D$ to obtain

$$B = (M+1)f, \text{ and } D = (N+1)g, \text{ or } D = (1/g - 1/C)^{-1}. \tag{3.2}$$

Substituting $E$ and $B$ for $C$, we obtain

$$D = \left[\frac{1}{g} - \frac{1}{(E-B)}\right]^{-1} = \left[\frac{1}{g} - \frac{1}{D_0/N_0 + (M_0 - M)f}\right]^{-1}. \tag{3.3}$$

This establishes the desired relationship between $A$ and $D$ in terms of the magnification $M$. The above expression for $D$ shows that, as the object moves to infinity, the total movement in the primary focal plane is $M_0 f$. The total movement in the secondary focal plane is further reduced from this value, which results in a wider range of focus over conventional imaging. In our example, the movement in the primary focal plane is $100\mu$m for an object distance of 1m to infinity. This translates into a mere $1.5\mu$m deviation in $D$. The magnification factor $N$ varies from 1/4 to 1/16. It is clear that, even with wide local apertures, the system is adequate for measuring depth while maintaining focus. Note that although objects remain in focus, the effective spatial resolution is decreased due to demagnification.

### 3.2.3 Depth Extraction

To obtain an expression for depth, consider Figure 3.9b and let $L$ be the distance between a pair of apertures and $\Delta$ be the displacement of the image between apertures. We estimate the distance $A$ from $\Delta$. Again, $E$ is adjusted to meet the desired magnification $N_0$ according to the other fixed parameters. The geometry of the configuration from the sensor to the primary focal plane gives:

$$C/L = D_0/\Delta. \tag{3.4}$$

Using the lens law for $A$ as a function of $B$ and making the substitution $B = E - C = B_0 + C_0 - C$, we obtain

$$A = \left(\frac{1}{f} - \frac{1}{B}\right)^{-1} = \left(\frac{1}{f} - \frac{1}{B_0 + C_0 - C}\right)^{-1}. \tag{3.5}$$

Solving for $A$ in terms of $\Delta$ gives the depth equation

$$A = \left[\frac{1}{f} - \frac{1}{(M_0 + 1)f + D_0/N_0 - D_0 L/\Delta}\right]^{-1}. \tag{3.6}$$

A characteristic of this sensor is that the amount of depth information available is a strong function of the object distance (the closer the object, the higher the depth resolution). We can quantify this by solving for $\Delta$ in terms of $M$, which gives

$$\Delta = \frac{D_0 L}{(M_0 - M)f + D_0/N_0}. \tag{3.7}$$

As $M$ increases, $\Delta$ rapidly approaches its limit of $D_0 L/(M_0 f + D_0/N_0)$.

The rate of change in $\Delta$ with $A$, i.e., $\partial\Delta/\partial A$, can be computed as a function of $\partial B/\partial A$

and $\partial\Delta/\partial C$. Setting $\partial C = -\partial B$ at the focal plane, it can be shown that

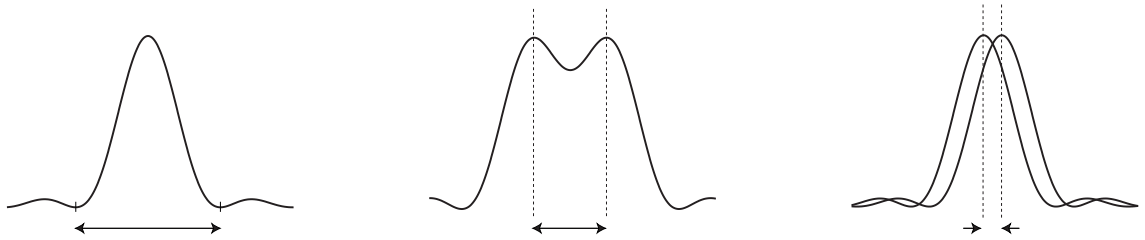
$$\partial\Delta/\partial A \approx -\frac{f^2}{A^2}\frac{DL}{C^2} \longrightarrow \partial\Delta/\partial A \approx -M^2N^2\frac{L}{D}. \tag{3.8}$$

For example, if assuming 0.5$\mu$m pixel pitch, the displacement between apertures can be estimated to within 0.5$\mu$m resolution. Further, assuming $L/D = 2$, the incremental depth resolution $\partial A$ is approximately 4cm at $A_0 = 1$m and 4mm at $A_0 = 10$cm. Decreasing pixel size allows for more accuracy in $\partial\Delta$ leading to higher depth resolution. Given sufficient signal-to-noise ratio at each pixel location, it is possible to determine feature location to dimensions smaller than the pixel size.

### 3.2.4 Spatial Resolution and Pixel Size

Clearly, the spatial resolution of our system is limited to the total number of pixels $mnk^2$. However, to establish overlapping fields of view, we set the magnification factor of the local optics to $N < 1$. Since each pixel is projected up to the focal plane by a factor of $1/N$, spatial resolution is reduced by $1/N^2$. Thus, the total available resolution is $\approx mnk^2N^2$. In our example, we assume a 16×16 array of 0.5$\mu$m pixels with a magnification factor of $N_0 = 1/4$. Thus, the maximum resolution is 16 times greater than the aperture count itself but 16 times lower than the total number of pixels.

The actual spatial resolution is limited by optical aberrations and ultimately by diffraction. The minimum spot size $W$ for a diffraction limited system is $\approx \lambda/\text{NA}$ (see Figure 3.10a), where NA= $n_i \sin\theta$ is the numerical aperture of the local optics, $n_i$ is the index of refraction of the dielectric and $\theta$ is the angle between the chief and the marginal rays. Using the Rayleigh criterion, the minimum useful pixel pitch is commonly assumed to be half the spot size (see Figure 3.10b). Assuming $n_i \approx 1.5$ in the dielectric stack, NA can be about 0.5, which gives a spot size of $\approx 1\mu$m. Thus, scaling the pixel beyond 0.5$\mu$m does not increase spatial resolution according to this criteria. Although no further increase in spatial resolution is feasible beyond the diffraction limit, depth resolution continues to improve as long as there are features with sufficiently low spatial frequencies. Indeed, the disparity between apertures can be measured at smaller dimensions than set by the diffraction limit as illustrated in Figure 3.10c.



(a) $W \approx \lambda/\text{NA}$, Diffraction limited

(b) $W \approx \lambda/2\text{NA}$, Rayleigh Criterion

(c) $W < \lambda/\text{NA}$, Multi-aperture

Figure 3.10: Spot size comparison.

## 3.3 Problems with Image Formation

A fundamental performance limitation in optical systems used for image formation is caused by diffraction. In Appendix A, there is a discussion of pinhole imaging systems in which diffraction and geometry are optimized to produce the smallest spot size possible. A pinhole camera is severely limited in both resolution and sensitivity. By using a refractive lens, a larger portion of light can be collected at a single photosite. The size of the aperture and the focal length are what determine the maximum resolution available in the system. In practice, lenses suffer from other aberrations which are discussed below.

### 3.3.1 Chromatic Aberration

The refractive index of a material is frequency dependent so a single lens is incapable of achieving the same focus for the entire visible spectrum. The refractive index for blue is larger than that for red so the effect shown in Figure 3.11 is produced. The way to correct for this effect is to use two different materials. Usually, the positive lens is made of a material with low dispersion while the negative lens is made with a high dispersion. The negative lens is weaker than the positive lens so that the net power is still positive. Glass materials are typically used because there has not been as much variation in the plastics available today.

### 3.3.2 Spherical Aberration

Glass lenses are usually made with spherical surfaces for ease of manufacturing. The spherical surfaces cause rays which are at different distances from the optic axis to converge at different points. Rays near the edge of the lens will have a shorter focus than the paraxial

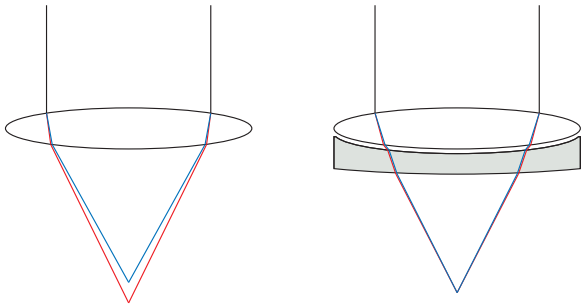


Figure 3.11: Chromatic aberration and correction.

region. The problem is monochromatic and occurs on and off-axis which results in image blur. A single lens can be corrected for spherical aberration by using the correct aspherical profile. Plastic lenses are often molded as aspherical to reduce this problem. Glass spherical lenses can be corrected with the use of symmetric doublets.

### 3.3.3 Distortion

Lens imperfections may cause distortion in the image. However, systematic distortion from the geometry of the lens may also occur. The image may appear to swell in the center like *barrel* distortion or sink in the center like *pincushion* distortion. Usually the amount of distortion is given as a percentage. Distortion can be minimized by using symmetric doublets.

### 3.3.4 Coma

Coma is the result of an aberration which causes an off-axis point to be blurred away from the optical axis. The effect is monochromatic and is usually described off-axis only. The problem may be reduced by ensuring that lenses that are symmetric about the aperture stop.

### 3.3.5 Astigmatism

Astigmatism is a result of different lens curvatures occurring across different planes. A point source will be spread unevenly, resulting in image blur. Oblique astigmatism occurs for off-axis rays with spherically ground lenses where there is a difference in curvature between the horizontal and vertical planes.

### 3.3.6 Field Curvature

Field curvature occurs off-axis where the outer region of the image is focused closer to the lens than the center region. A planar object gets projected as a curved image. It is essentially a difference in focal length for rays originating at a large angle from the optical axis. A lens with field curvature could be stopped down to maintain a larger depth of field.

## 3.4 Lens Correction with Multi-Aperture system

There are several advantages of the multi-aperture array compared to the focal plane array. In the focal plane array, each pixel has a color filter above it that is completely surrounded by pixels with filters of a different color. Color interpolation must be applied to represent a full color image. Aliasing and color crosstalk are issues that must be dealt with. In addition, the optics of the system must be corrected for chromatic aberration. In the multi-aperture array, each module may contain its own color band. With this scheme, color crosstalk is eliminated and color interpolation is not needed. In addition it may be possible to use optics that are not corrected for chromatic aberration. Glass optics are usually needed to correct for this aberration and at least two elements are used. Figure 3.12 shows hows a single lens produces separate focal planes dependent on color.

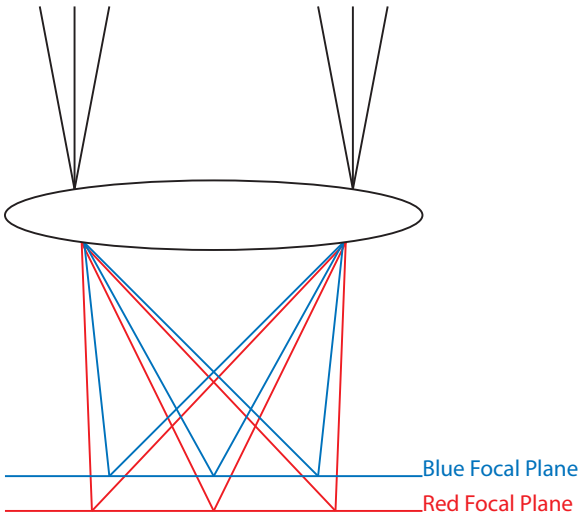



Figure 3.12: Color planes formed by chromatic aberration.

Since this method of imaging requires each module to focus into the objective focal plane independently, it is possible to allow each color module to focus independently of the

others. Figure 3.12 shows how each color module focuses into its own color plane.

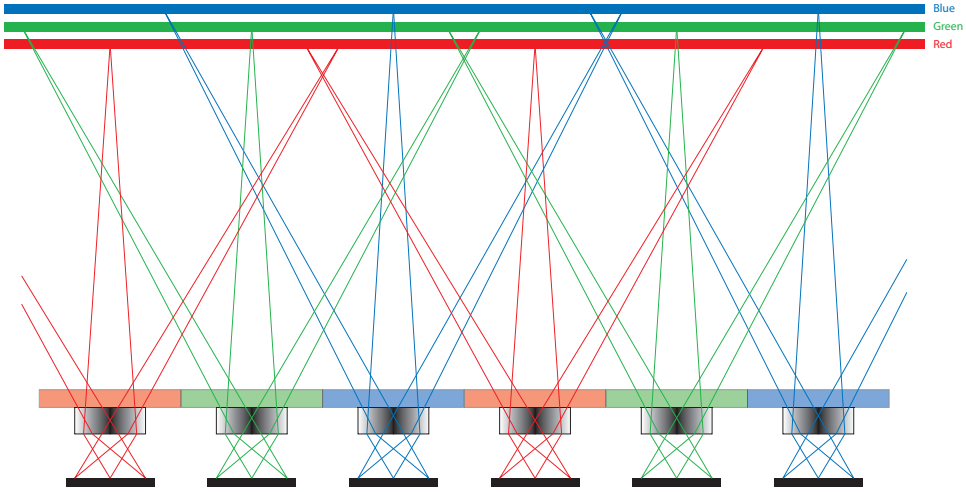



Figure 3.13: Color channel separation by focal plane.

While chromatic aberration can be solved by customizing the local optics for each color channel, other optical aberrations can be corrected or improved by designing the local optics in conjunction with the objective. A simple example is in the correction of field curvature as shown in Figure 3.14. Each aperture's focal length becomes a function of its position in order to bring all objects into focus. Computation is then required to adjust for any variation due to the local magnification of each subarray.

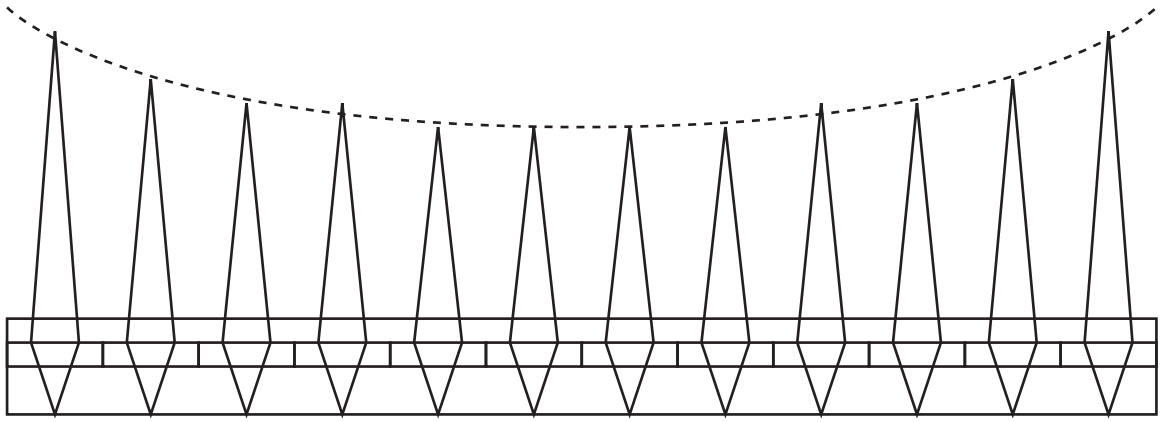


Figure 3.14: Correction for field curvature with multi-aperture array.

Multi-aperture imaging with a objective lens has the potential to reduce the optical requirements required by focal plane arrays and yet produce more robust systems that may

have additional benefits because of the flexibility given at the device plane. One could imagine a single external lens made of plastic. This lens can be aspheric to eliminate the spherical aberration. One of the main driving forces behind glass optics is the ability to correct for chromatic aberration. Once in glass, several lenses are then added into the design because of the need to correct for the other aberrations. Lenses designed for low F numbers are fairly complicated and systems that use them have poor depth of focus. It is possible with the distributed array to use low $f/\#$ optics to achieve high sensitivity while still maintaining depth of focus across an image as long as each module can be independently controlled. An $f/1$ system with all objects in focus is conceivable.

### 3.4.1 Multi-Aperture Array for Close Proximity

A multi-aperture array may be used without external optics in the near field. Instead of imaging the focal plane of a objective lens, it is possible to image the object directly. Section 3.6 describes a problem in which the redundancy, and hence, the resolution of such systems is dependent on the object distance. Nevertheless, for some fixed distance applications such as optical mice and finger print scans, this system has the advantage of eliminating external optics. Figure 3.15 shows how each array contains a field of view that overlaps with the next. The maximum field of view of this system is essentially the same size as the global array.

In order to extend close proximity imaging to greater object distances with a wider field of view, each module may be physically tilted in the appropriate direction. This is similar to the concept of the insect eye except that each module is an array instead of a single facet. Instead of tilting the module, which would not be easy in a semiconductor, a prism may be used. Each module would need a different prism that increases in slope as the array progresses to the edge. This redirects the field of view of each module to extend the entire field of view of the system. Instead of using prisms, a concave lens can be placed on top of the distributed array as shown in Figure 3.16.

The concave lens is not an imaging lens so the focal planes remain at the individual modules. Also the lens does not need to be placed with any separation at the interface which allows the stack height to be short. Since there is no focal plane produced by the lens, the system unfortunately suffers in sensitivity. Only the local modules and the surrounding neighbors with redundancy collect light from the source. This is different from the distributed array with an objective lens where the objective lens focuses the source from a much larger collection region.

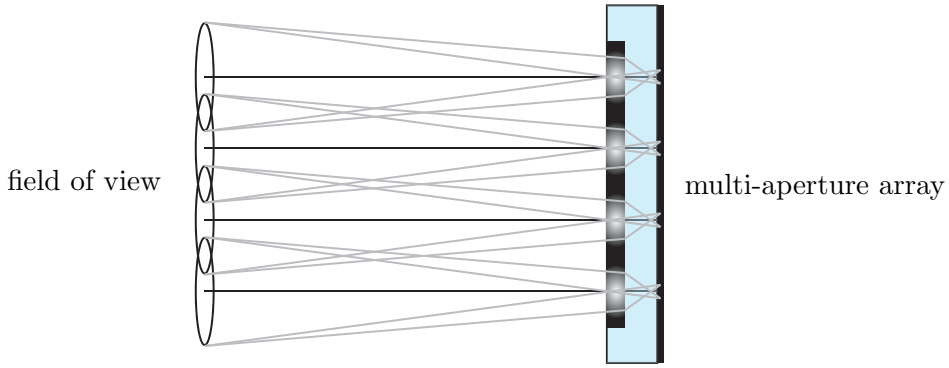



Figure 3.15: Multi-aperture array without external optics.

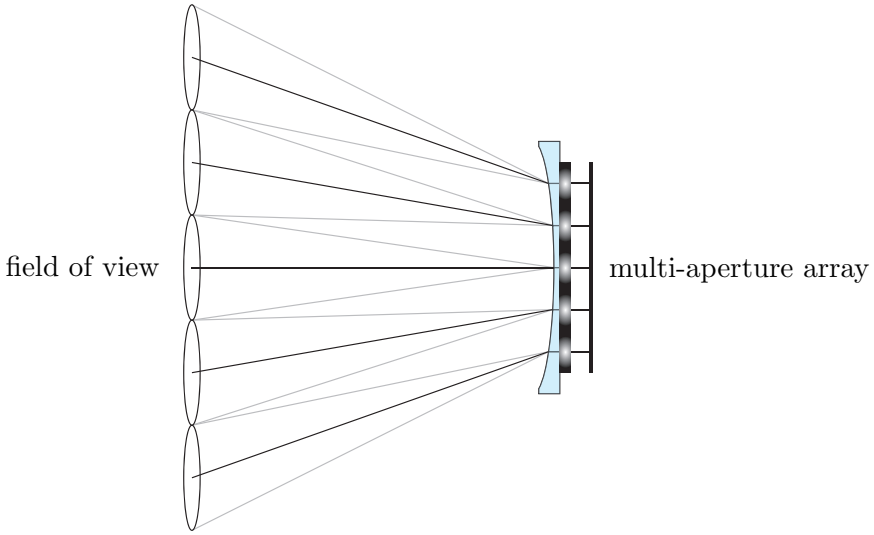



Figure 3.16: Multi-aperture array with concave lens.

## 3.5 Local Optics

Microlenses have been a part of focal plane imaging for over a decade. These devices are used as concentrators in order to achieve higher optical fill factors and in some circumstances reduce crosstalk. These devices may also be used as imaging lenses as long as the aberrations are kept to a minimum. Suppose that each module were to have its own color channel. Each microlens would need a slightly different focal length in order to compensate for the chromatic aberration in the external lens as well as the chromatic difference in focal length due to color channel separation. Aberration needs to be controlled and a mechanism for

tuning the focal length across color channels is necessary for a given resolution target[1]. In order to achieve reasonable resolution over a small area, pixels should be smaller than pixels in conventional focal plane arrays because some resolution will be lost due to the redundancy between modules. Suppose that pixels are created near the size of $\lambda$. This will require fast, diffraction-limited optics. In order to resolve to $\lambda$, the $f/\#$ corresponding to the local optics must be $f/1$. While it is difficult to achieve diffraction-limited performance at this wide of an aperture, the small scale of the lens makes it much easier to control the geometric aberrations. If local modules can have a relatively long focal length compared to their format size, the field of view is more narrow and local optics may be easier to design for a given resolution requirement.

An approach to tuning the focal lengths of the individual modules or by color channel is to use electrooptics. Appendix B gives an example of the electrooptic effect in Lithium Niobate. By applying an electric field to the dielectric region just behind the microlens it is possible to increase or decrease the refractive index. This changes the back focal length of the lens. With this technique, variable focal plane imaging may be possible.

It is also possible to create a lens by applying a gradient electric field in the dielectric region. A lens changes the phase of the incident wave as a function of the radial distance to the optical axis. A graded index may be created such that it produces the correct phase alteration to bring an image to a focus. It seems possible to create a lens in one dimension with this approach. This would emulate a cylindrical lens which brings a point source to a focused line. Two cylindrical lenses rotated by 90 degrees will function like a spherical lens where the point source is brought to focus at a point instead of a line. With significant separation between the lenses, each must have a different focal length which will distort the image. Using several stacks of cylindrical lenses may reduce this effect.

## 3.6 Resolution and Redundancy

In order to calculate the recoverable resolution from a multi-aperture image sensor, it is necessary to account for the amount of redundant data that each imager module produces. Obviously, if each imager module images the same source, all data would be redundant and the total resolution would be equal to the resolution of a single imager module. On the other extreme, if each module images non-overlapping regions of the source, the total resolution would be equal to the total number of photosites in the multi-aperture array. To

[1] A system of complex external lenses could alleviate this requirement.

achieve high resolution, the system should have as little redundancy as possible. However, high redundancy is desirable for robustness. One of the advantages of distributed imaging systems is that redundancy is an adjustable parameter. This has advantages for color processing and sensitivity.

### 3.6.1 The Parameter $\mathcal{R}$

We define a parameter $\mathcal{R}$ such that it can be used in calculations for resolution. In practice, it will be more useful to work with the non-redundancy $(1 - \mathcal{R})$ since it is easier to picture an increase in resolution as being attributed to the non-redundant portion of the image. Figure 3.17 shows the three overlapping fields of view created by each module. The shaded region shows the area that has redundancy.

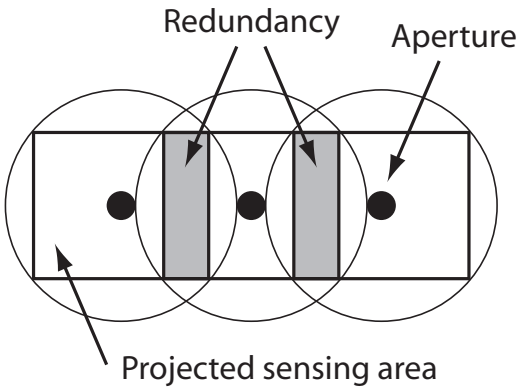



Figure 3.17: Overlapping fields of view showing redundancy

Using Figure 3.18 we define the redundancy for a single module belonging to a larger 3×3 array of modules. The module labeled number 5 is completely overlapped by the other neighboring modules. If this module were to become defective, the image that it is normally sampled at this location can still be recovered because it is fully redundant (in field of view). However, we define the redundancy for module 5 as only the redundancy with respect to the next module in the x-direction and the next module in the y-direction as indicated in the figure. The total redundancy is given by the equation:

$$(1 - \mathcal{R}) = (1 - \mathcal{R}_x)(1 - \mathcal{R}_y)$$

In the example in Figure 3.18, $\mathcal{R}_x = 0.5$ and $\mathcal{R}_y = 0.5$ which gives a non-redundancy of 0.25. The value of $\mathcal{R}$ in this case is 0.75.

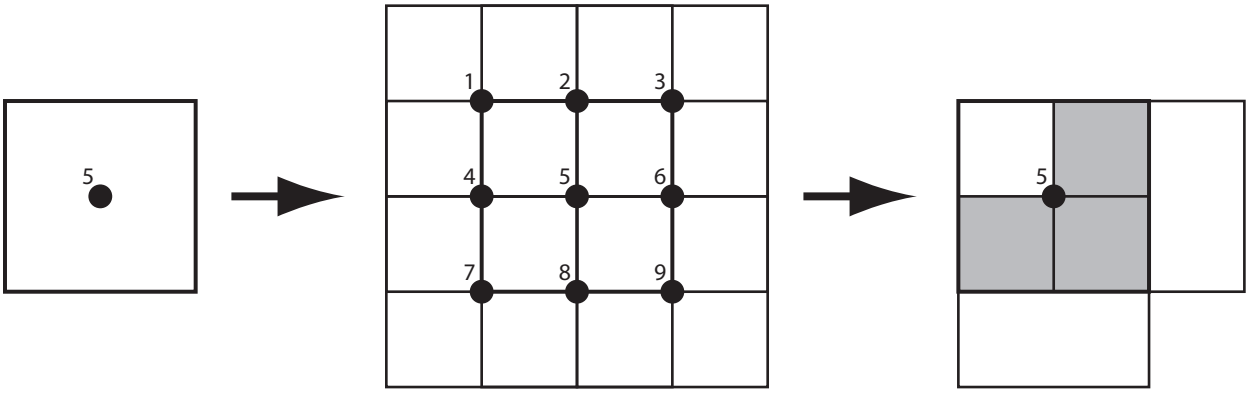



Figure 3.18: Illustration for the calculation of $\mathcal{R}$ .

### 3.6.2 Calculation of Resolution

Using the definition of $\mathcal{R}$ , we can calculate the maximum resolution that a system may achieve. As a first pass, we can estimate the resolution to be approximately equal to the total number of pixels in the system multiplied by the non-redundancy. A more rigorous analysis follows:

$$\begin{aligned} \text{Let } i &= \text{The horizontal pixel count of the repeated module} \\ \text{Let } j &= \text{The vertical pixel count of the repeated module} \\ \text{Let } n &= \text{The number of times the module is repeated horizontally} \\ \text{Let } m &= \text{The number of times the module is repeated vertically,} \end{aligned}$$

$$\text{Resolution} = ij[1 + (n-1)(1-\mathcal{R}_x) + (m-1)(1-\mathcal{R}_y) + (n-1)(m-1)(1-\mathcal{R}_x)(1-\mathcal{R}_y)]$$

Since we are only interested in parts of the array that are consistent, we would choose to neglect the borders of the array where there is no redundancy. We are specifically interested in the area of the array in which there is both redundancy in $x$ and $y$. This reduces the equation for resolution to:

$$\text{Resolution} = (ij)(n-1)(m-1)(1-\mathcal{R}) \tag{3.9}$$

This is the essentially the result we expect from the first pass. Figure 3.19 demonstrates a simple case for calculating the resolution. The shaded region indicates the region of consistent overlap and hence, the region in which the resolution is of interest. There are 4×4 modules in the array, each with resolution $ij$ and $\mathcal{R} = 0.75$. The result is a resolution of $(9/4)ij$. This can be confirmed by counting the effective number of modules covered by the shaded region. There are nine quarter-sized modules of resolution $ij$.

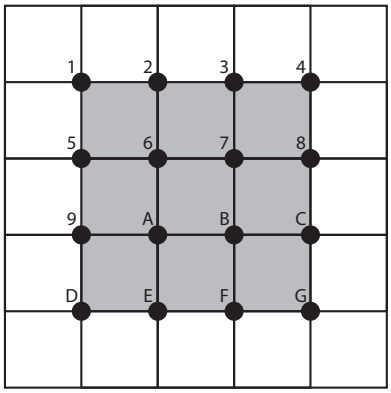



Figure 3.19: Example for calculation of resolution.

### 3.6.3 Choosing $\mathcal{R}$ for Color Recovery

There is an advantage of having redundancy between modules when color imaging is required. A typical focal plane array consists of a mosaic array of color filters where each photosite contains a red, green or blue filter[2]. The data is then interpolated to estimate values for the two missing colors at each photosite. If an image is well focused or additional gratings are not used, color aliasing occurs in such systems. Often 3-CCD cameras are employed for professional use to eliminate this problem. In multi-aperture arrays, each module may contain a single color channel as indicated in Figure 3.20. With no redundancy between modules, the individual color planes contain missing data as indicated in the figure. Interpolation is then required, which leads to artifacts. Figure 3.21 shows that, with the right amount of redundancy, consistent color planes may be produced which eliminate color aliasing. For color imaging systems, a minimum redundancy of 75% is necessary to achieve this result.

[2]Complimentary color filters are also used but rarely have an advantage over RGB.

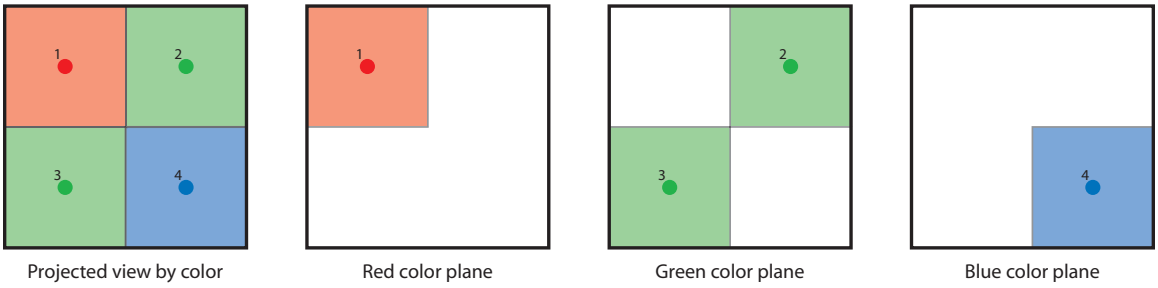



Figure 3.20: Modules with separate color channels and $\mathcal{R} = 0$.

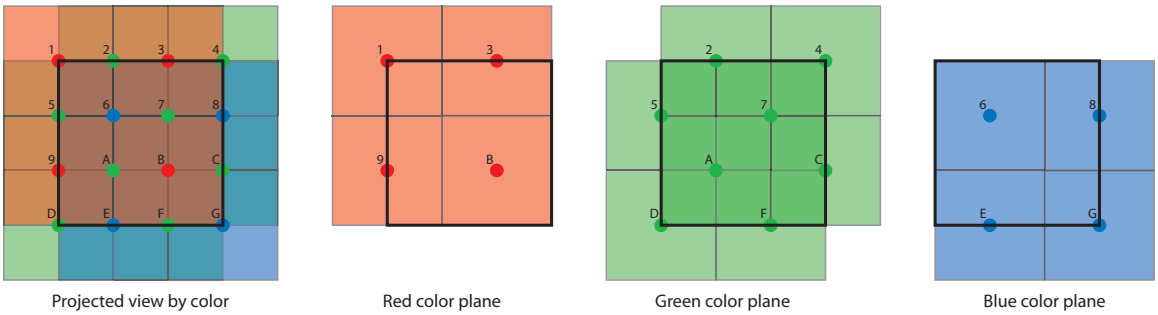


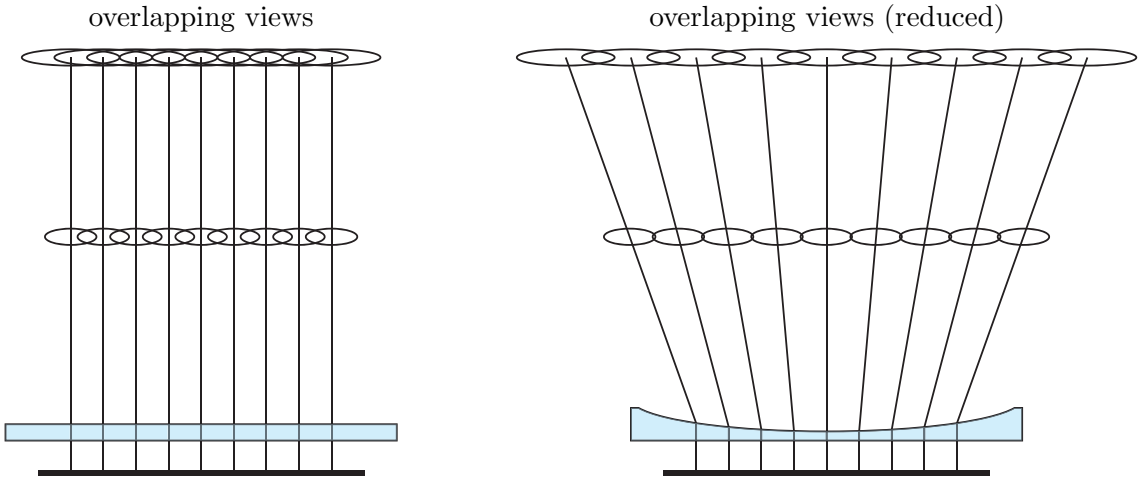

Figure 3.21: Modules with separate color channels and $\mathcal{R} = 0.75$.

### 3.6.4 How $\mathcal{R}$ Varies with Object Distance

Without any external optics, the distributed modules image the source directly. With a lens in front of the multi-aperture image sensor, the modules re-image the real image of the object in the focal plane of the objective lens. In the case without external optics, $\mathcal{R}$ is a strong function of object distance. As the object distance approaches infinity, $\mathcal{R}$ approaches 1. The resolution of the system at that point is equivalent to that of a single module. If the object becomes too close to the sensor, the image becomes under-sampled. One of the purposes of using an objective lens is to control $\mathcal{R}$ . Even if a lens with a small $f/\#$ is used, the distance from the sensor to the focal plane of the lens varies over a small amount as the object distance changes. Since there is only minor movement in the focal plane, the value of $\mathcal{R}$ remains fairly constant. This allows for a wide range of object distances with minor effect on the final resolution of the image.

Figure 3.22a shows how the redundancy grows as the object distance increases if an objective lens is not used. This makes such a system useful only over a certain range. Figure 3.22b shows how this situation can be rectified with the addition of a concave lens. This lens is not used to form an image but rather to radially alter the field of view for

the modules of the distributed array. Essentially each module sees a prism with a pitch dependent on its position in the distributed array. By matching the field of view with the correct curvature of the lens, it is possible to image in the far field without significant variation in $\mathcal{R}$ .




(a) Far-field imaging (b) Optical correction to reduce $\mathcal{R}$ .

Figure 3.22: Reduction in $\mathcal{R}$ by optical correction.

## 3.7 Sensor Architectures

The imager module architecture used in repetition for a multi-aperture image sensor may contain a wide variety of components. This section explores the different levels of components that may be added to a module and what applications they may serve. This ability to configure a module is shown to be one of the advantages of using the multi-aperture configuration, since we are no longer tasked with the design of a continuous array of pixels.

### 3.7.1 General Configuration

Every module to be discussed contains an active array where the image sensing physically occurs and a peripheral circuit section where the components for a specific performance requirement or configuration exist. Figure 3.23a shows the basic imager module while Figure 3.23b shows how the module may be used in an integrated array of sensors. In practice, the single module may be repeated across an array millions of times.

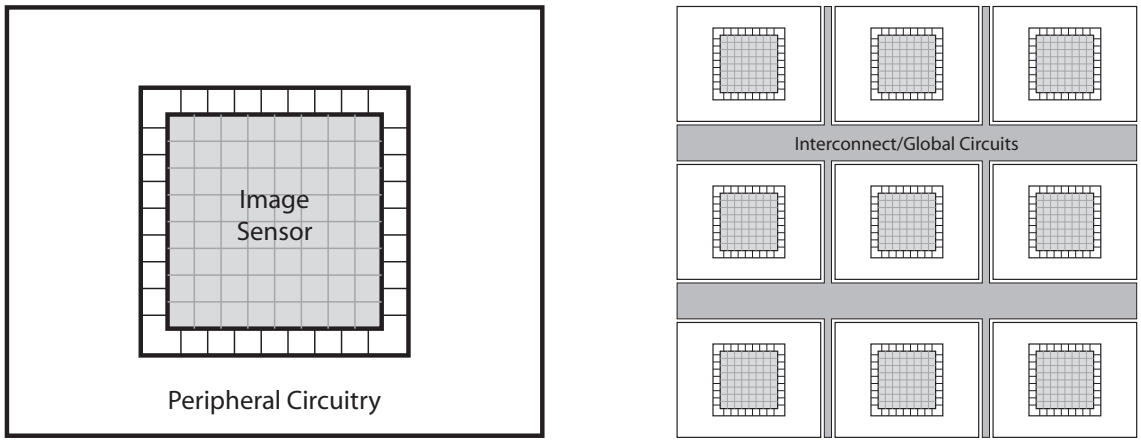



Figure 3.23: General image sensor module (a) and array (b).

### 3.7.2 Basic Configuration

It is desirable to achieve a large optical fill factor for both sensitivity and resolution. The redundancy factor $\mathcal{R}$ is what controls the trade off between sensitivity and resolution at a given fill factor. The most basic configuration should yield the largest fill factor. Figure 3.24 shows an architecture which includes an imaging array, a readout section and an amplifier. The array is readout by one pixel at a time and buffered with the amplifier so that the analog value may transferred to another section of the chip for readout or conversion. The approach is similar to an APS architecture except that the concept of a regular array with focal plane image formation does not apply. Therefore, there is flexibility in the way control signals are driven and in the design of the amplifier. The downside of this architecture is that each row or column of arrays must integrate on a shifted cycle in order to progressively readout each cell.

### 3.7.3 Configuration with Frame Buffer

By adding a frame buffer, every pixel and every imager module may integrate on the same cycle. Figure 3.25 shows an architecture which includes a frame transfer section identical to the active array except that no photons may interact with the area. This allows for full-frame capture of an image which, in certain applications, is a major advantage over progressive scan. At high resolution, it is difficult to capture the entire image in a respectable amount of time. This is why many high resolution still cameras still require a mechanical shutter.

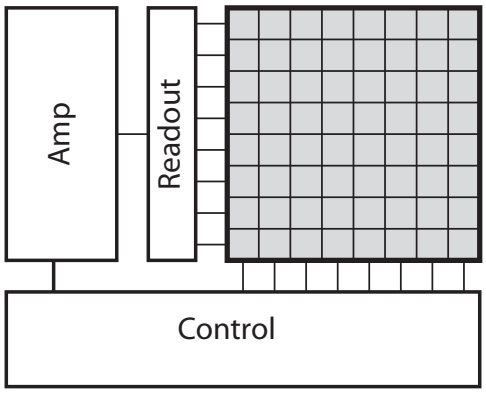



Figure 3.24: Module including control, readout and amplifier.

With this scheme, the camera may have arbitrarily high resolution and still maintain the same frame capture characteristics.

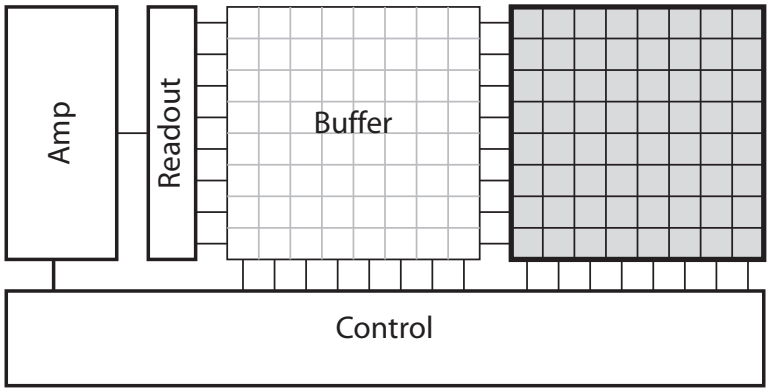



Figure 3.25: Module including frame buffer, control, readout and amplifier.

### 3.7.4 Configuration with ADC

There is not a fundamental advantage to bringing the ADC to the imager module shown in Figure 3.26, however it can be desirable for distribution of circuitry. Otherwise the array would contain a section of imager modules and then a section of readout or A/D converters. This is really a matter of choice and design capability. However, in the multi-aperture array, it is not undesirable to bring the ADC to the imager module as long as the number of modules remains constant. This is a difference between a focal plane array and the multi-aperture array. The multi-aperture array may easily distribute the optical, active area whereas the focal plane array must maintain regularity. Any additional circuitry at the

pixel level in a focal plane array takes away from the signal but as long as the redundancy and number of modules remains constant, the global fill factor in the multi-aperture array does not affect sensitivity or resolution.

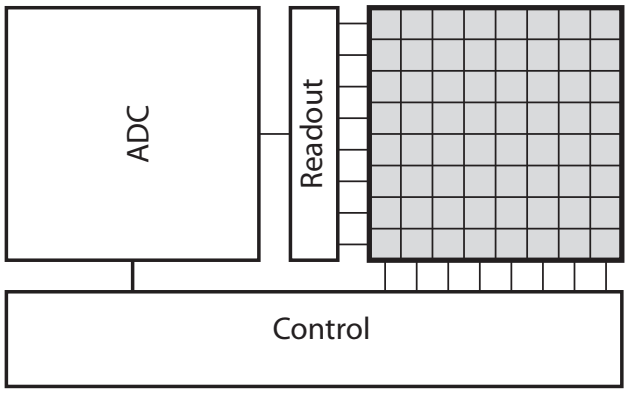



Figure 3.26: Module including control, readout and ADC.

### 3.7.5 Configuration with Frame Buffer and ADC

By adding both a frame buffer and the ADC to the module as shown in Figure 3.27, a near stand-alone imager module is created. Again, the addition of the ADC is useful for better distribution and easier repetition and configuration. The ADC could be shared between adjacent modules in order to save space. The amount of sharing would be dependent on the area, power and speed required for the targeted application. Again, flexibility is permitted.

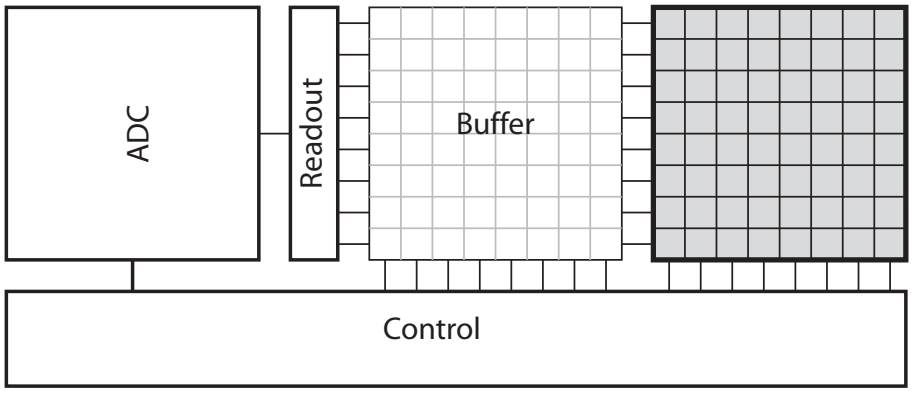



Figure 3.27: Module including frame buffer, control, readout and ADC.

### 3.7.6 Local Processing

It is possible to add some processing functionality to each module as shown in Figure 3.28. Since reconstruction of the image in the focal plane will likely involve repetitive operations over small areas, local processing can be advantageous. Color processing can be performed very efficiently as long as there is enough redundancy between modules. The processing could be done during readout so only a limited amount of memory would be needed.

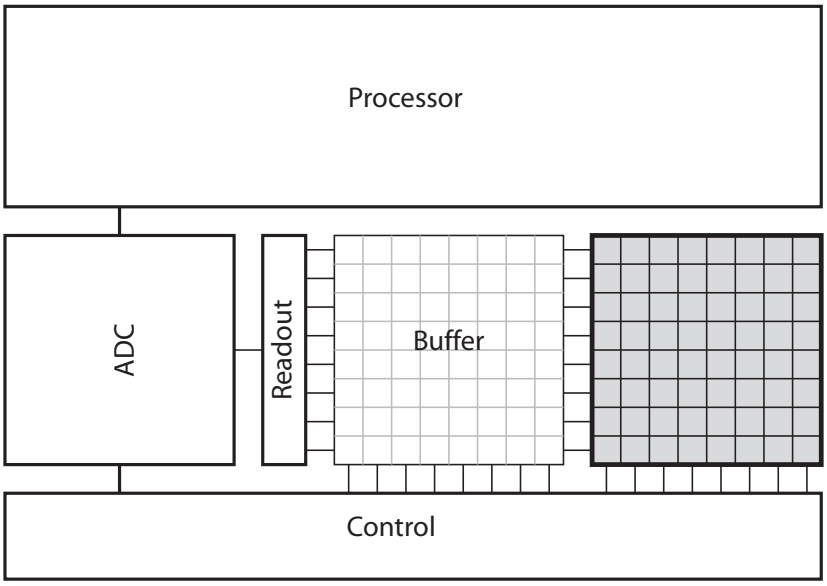



Figure 3.28: Module including frame buffer, control, readout, ADC and processor.

### 3.7.7 Addition of Memory

By adding memory to the module, an entire frame may be stored indefinitely. Although it is more area efficient to store the frame in a frame transfer buffer, digital memory may be needed if readout time is much longer than the lifetime of the frame transfer buffer. Figure 3.29 shows the addition of memory. The size of the memory is likely to be large compared to the imager array. It is possible to reduce the memory size by storing pixel differences rather than pixel magnitudes. Since each module overlaps with the next because of redundancy, pixel differences may be used to recover the entire image. The amount of resolution needed for a given pixel pair would be a function of the maximum MTF at the highest frequency. The differences could be stored as variable length.

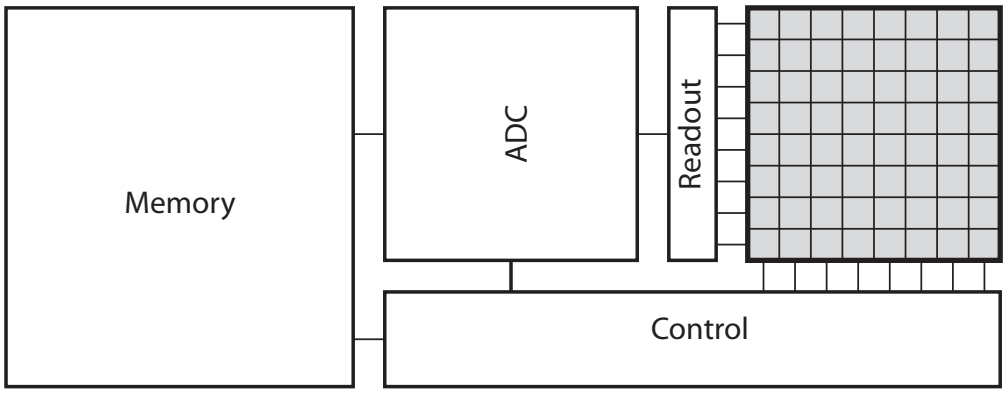



Figure 3.29: Module including control, readout, ADC and memory.

### 3.7.8 Addition of Memory and Processor

One of the disadvantages of using a frame transfer buffer instead of digital frame buffer is that the frame transfer buffer can not be easily accessed randomly. In addition, reading a CCD transfer buffer is destructive. With the addition of a processor, certain functionality may require multiple reads across modules. Adding a digital memory may be a better fit for the processor. This architecture is shown in Figure 3.30

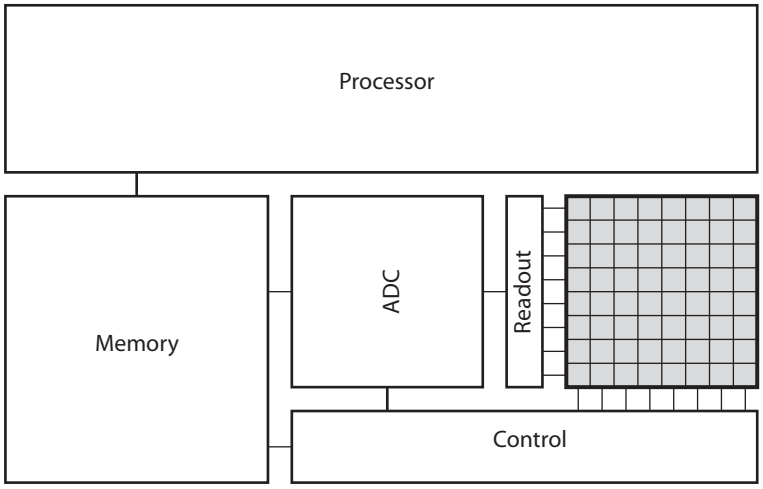



Figure 3.30: Module including control, readout, ADC, memory and processor.

### 3.7.9 Addition of Frame Buffer, Memory and Processor

Although a digital memory buffer may replace the frame transfer buffer, it may be important to keep a frame transfer buffer available for quick transfers. Otherwise, the data path involves the A/D converter and the readout speed would be limited by this design. Figure 3.31 shows a complete imager module including all of the components discussed in this chapter. There are still other components that can be added to the module depending on the application. Some of the specific details of the control block have been left vague. There are a number of options such as pixel binning and cycle transfer control that could be explored.

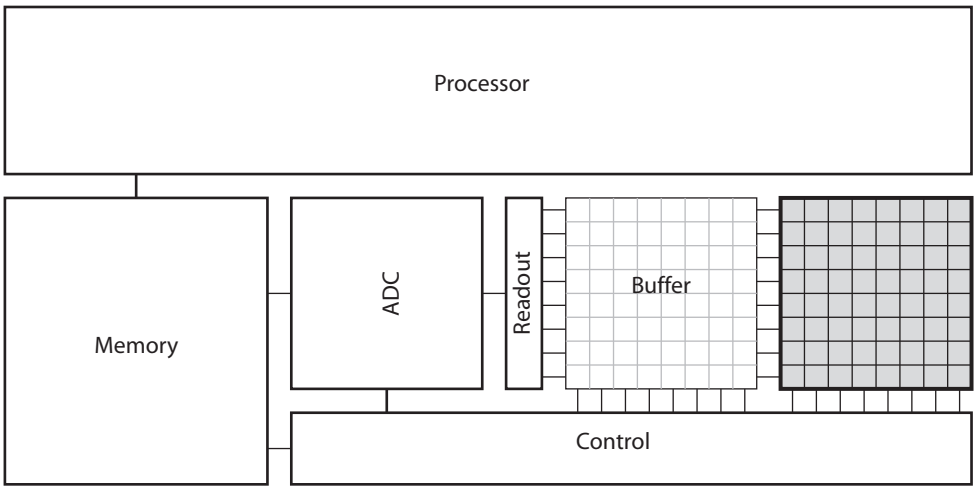



Figure 3.31: Module including frame buffer, control, readout, ADC, memory and processor.

# Chapter 4

# Submicron Pixel Design

Building submicron CMOS pixels with acceptable imaging performance at wide apertures is challenging in part due to the thick dielectric stack height and optical occlusions resulting from the use of metal layers in the pixel. It has recently become possible to implement deeply scaled CCD pixels in modern CMOS technology due to the narrow poly gap spacing afforded by device scaling. This eliminates the need for processing multi-layer polysilicon and allows for electrode widths roughly equal to the gap spacing, which increases the effective quantum efficiency. Since the poly gap region normally creates either a pocket or a barrier, we make use of this feature to achieve single electrode charge confinement. The need for anti-blooming and high charge transfer efficiency are relaxed by the multi-aperture architecture where smaller subarrays are distributed globally. Here, blooming is constrained within the aperture and the number of charge transfers is scaled proportionally by the total number of apertures. By relaxing some traditional imaging requirements and making use of an alternative architecture, this work produces pixels much smaller than the trend shown in Figure 4.1.

This chapter describes the design and characterization of 3 types of CCD structures implemented in 0.11$\mu$m CMOS technology: surface-channel, buried-channel, and pinned phase buried-channel. Each CCD structure differs in the location of charge storage during the integration time and in the sequencing of the electrodes during charge transfer. The surface-channel design has a large well capacity even at 0.5$\mu$m pixel pitch but the transport of charge near the Si-$SiO_2$ interface decreases the charge transfer efficiency (CTE) and increases dark current. The buried-channel CCD design attempts to move charge confinement away from the surface to improve CTE. A pinned phase buried-channel design is implemented to improve both dark current and charge transfer efficiency by shielding

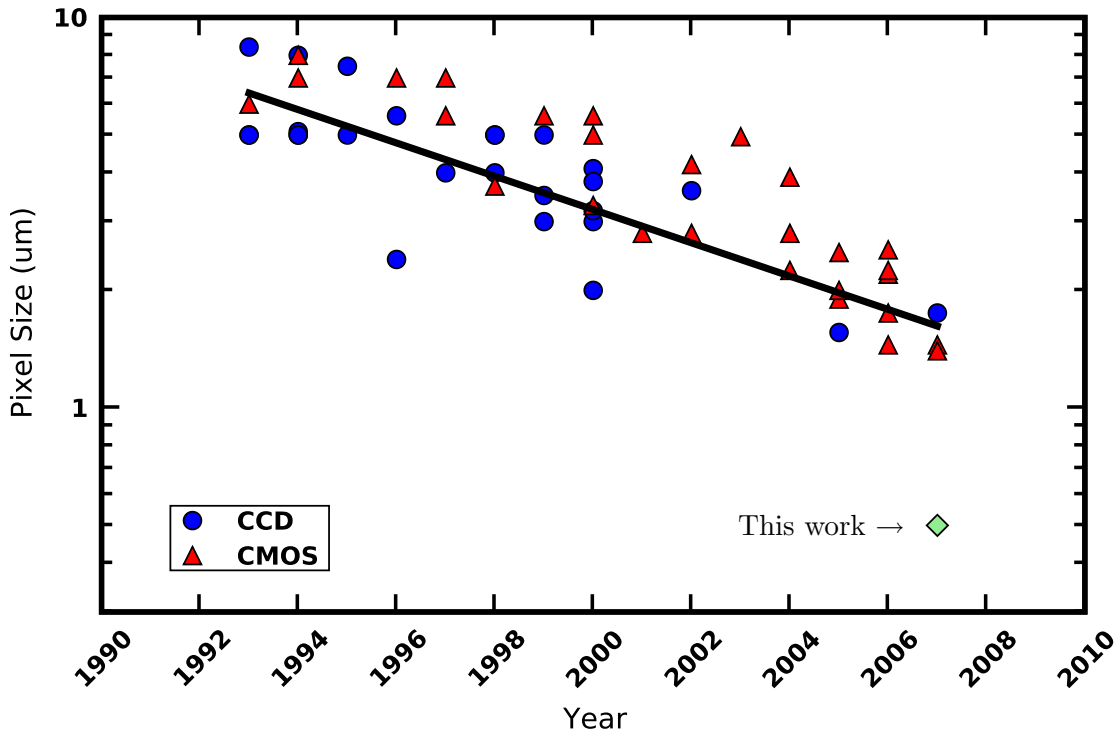



Figure 4.1: Recent pixel scaling trend including this work.

dark current generation at the surface from the confinement region. The poly gap space is used as an area for self-aligned implantation to create barriers that allow for inversion of the channel during the integration period. To characterize the designs, test structures were fabricated comprising arrays of 16×16, pixel FT-CCDs, each with source follower readout. The design, fabrication, and operation of the image sensor will be discussed first, followed by the simulation and characterization results.

## 4.1 Design

An FT-CCD architecture is used to minimize pixel pitch and to eliminate metal layers in the active imaging area. The format of each sensor is approximately equal to the dielectric stack height, which enables $f/1$ scale apertures for high sensitivity. We take advantage of the fact that the image capture is a distributed process by partitioning the subarrays into regions of high fill factor and occlusion free optical paths. Implementing the CCDs in CMOS technology enables fast multi-array readout along with the integration of analog and digital circuits. Each test structure consists of a pixel array, a storage array, a horizontal (H)-CCD, and a source follower readout circuit (see Figure 4.2). The storage array is covered by metal layers that are also used to distribute global control lines (see Figure 4.3a). A photomicrograph is shown in Figure 4.3b and an SEM image is shown in Figure 4.4. Figure 4.5 shows a 4×4 section of the pixel array with critical dimensions labeled. For the 0.5$\mu$m designs, the electrode width is 320nm with a poly gap width of 180nm. The channel stop with is 250nm for the surface-channel device. The buried channel devices use a P-type stopper implant of approximately the same mask-patterned width. Figure 4.6 shows SEM cross-sections in both directions of the surface-channel device. The design process for each device is briefly explained in the sections that follow.

### 4.1.1 Surface-channel CCD

The surface-channel device makes use of common structures available in standard CMOS technology to achieve single electrode charge confinement. In order to create a potential well for the collection of photo-generated electrons, it is necessary to bias the electrodes relative to the substrate such that free carriers in the substrate are pushed away from the surface. Since the bulk semiconductor is doped with boron, negative fixed charge is left in the depletion region to compensate for the positive charge on the electrode. Therefore, the maximum potential in the semiconductor bulk occurs at the surface as shown in the

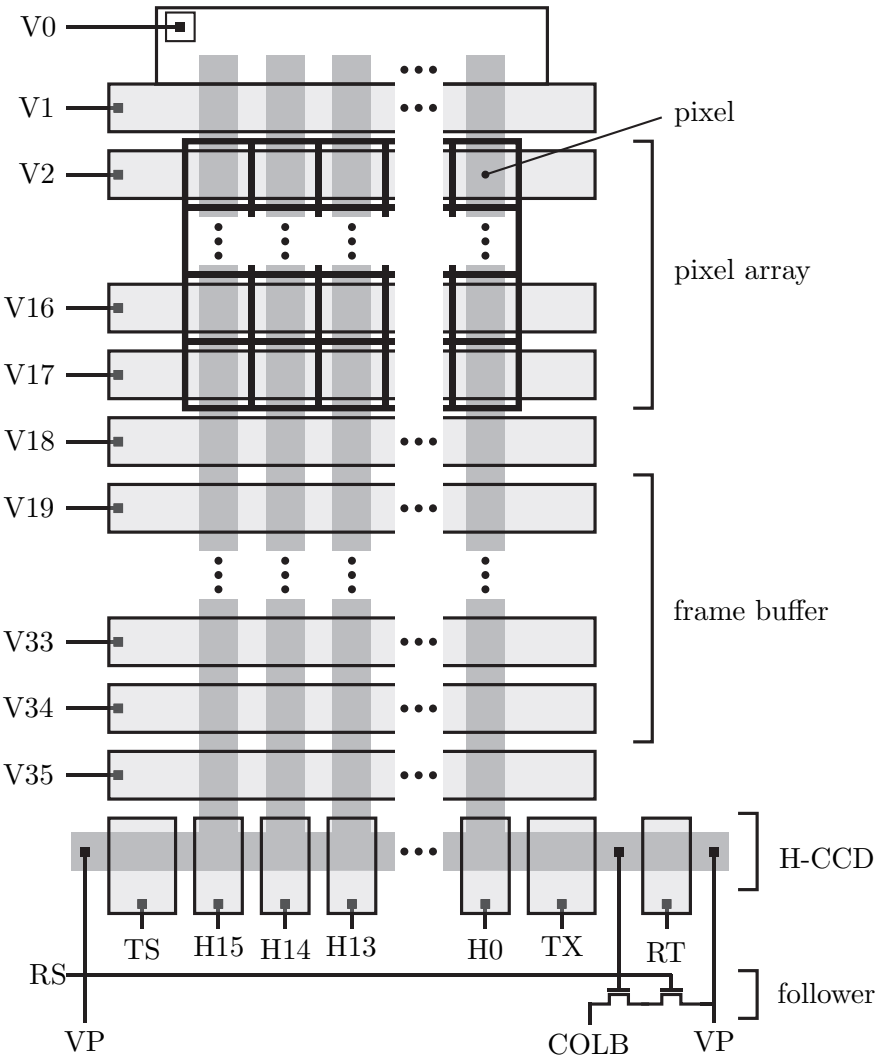
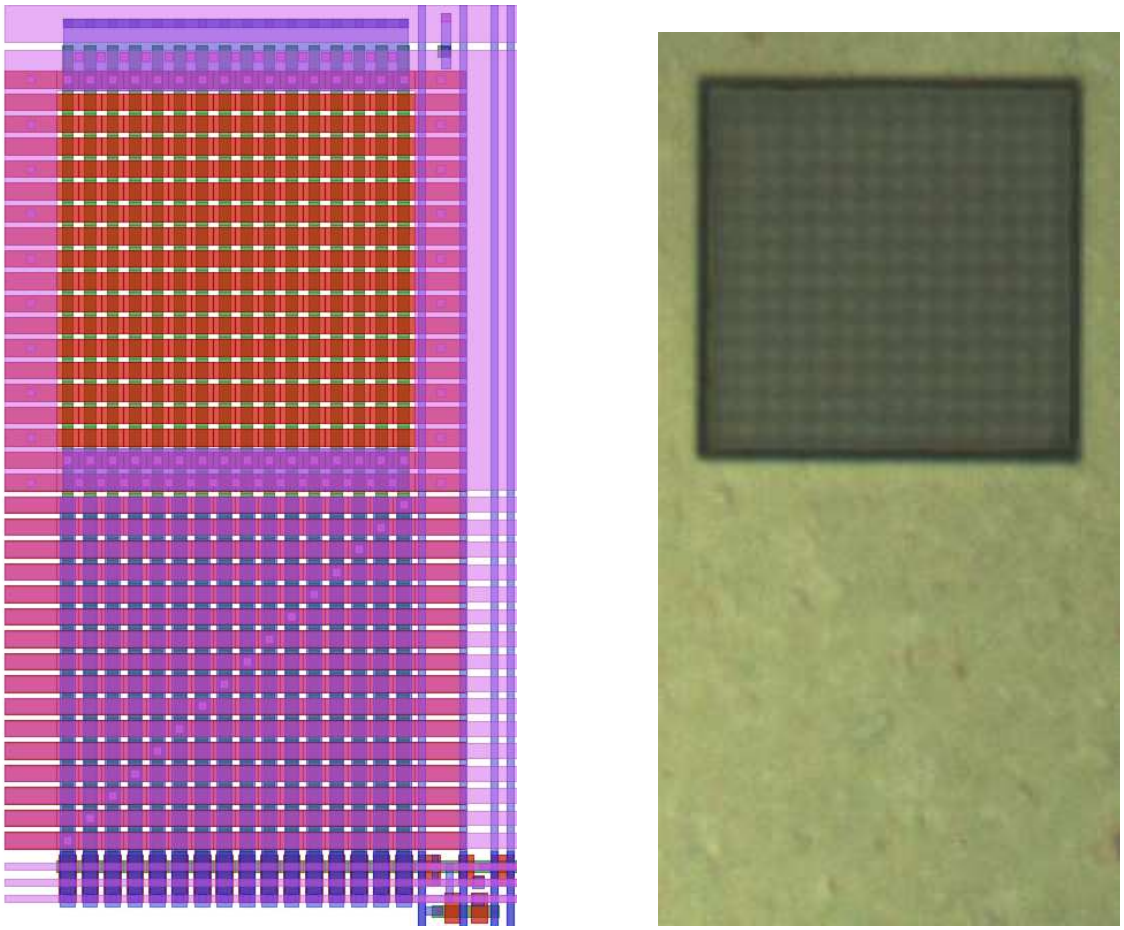



Figure 4.2: FT-CCD schematic showing the pixel array, frame buffer, H-CCD and follower readout.



(a) Mask Layout (b) Photomicrograph

Figure 4.3: Mask layout and photomicrograph of a fabricated 16×16 FT-CCD. Two images in the photomicrograph are combined to simultaneously focus on the pixels and the top metal.

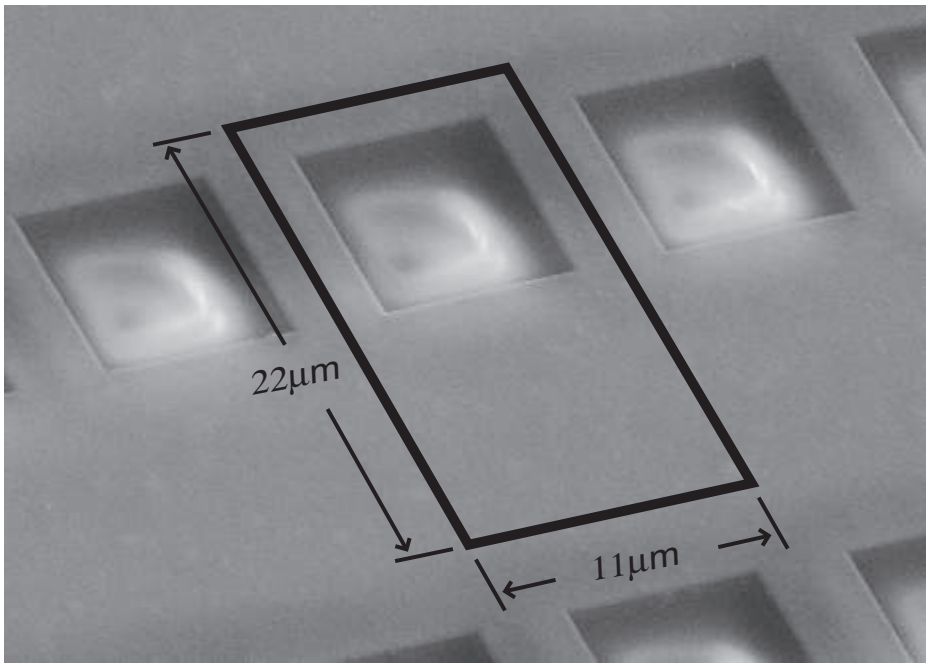



Figure 4.4: A view of the fabricated test structures showing the metal opening on top of the 16×16 image sensor active area.

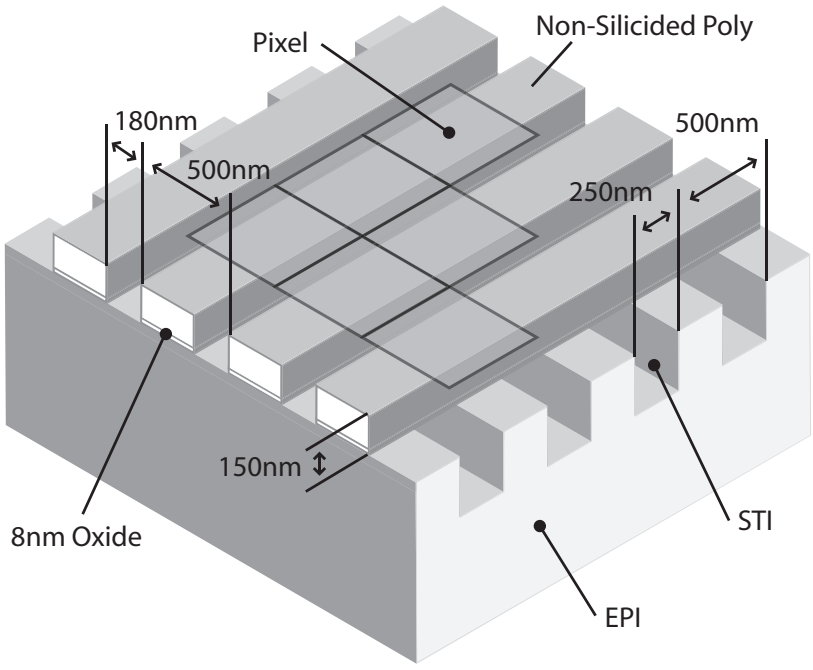



Figure 4.5: A view of a 4×4 portion of the pixel array showing pattern-doped poly and STI channel stops and critical device dimensions.

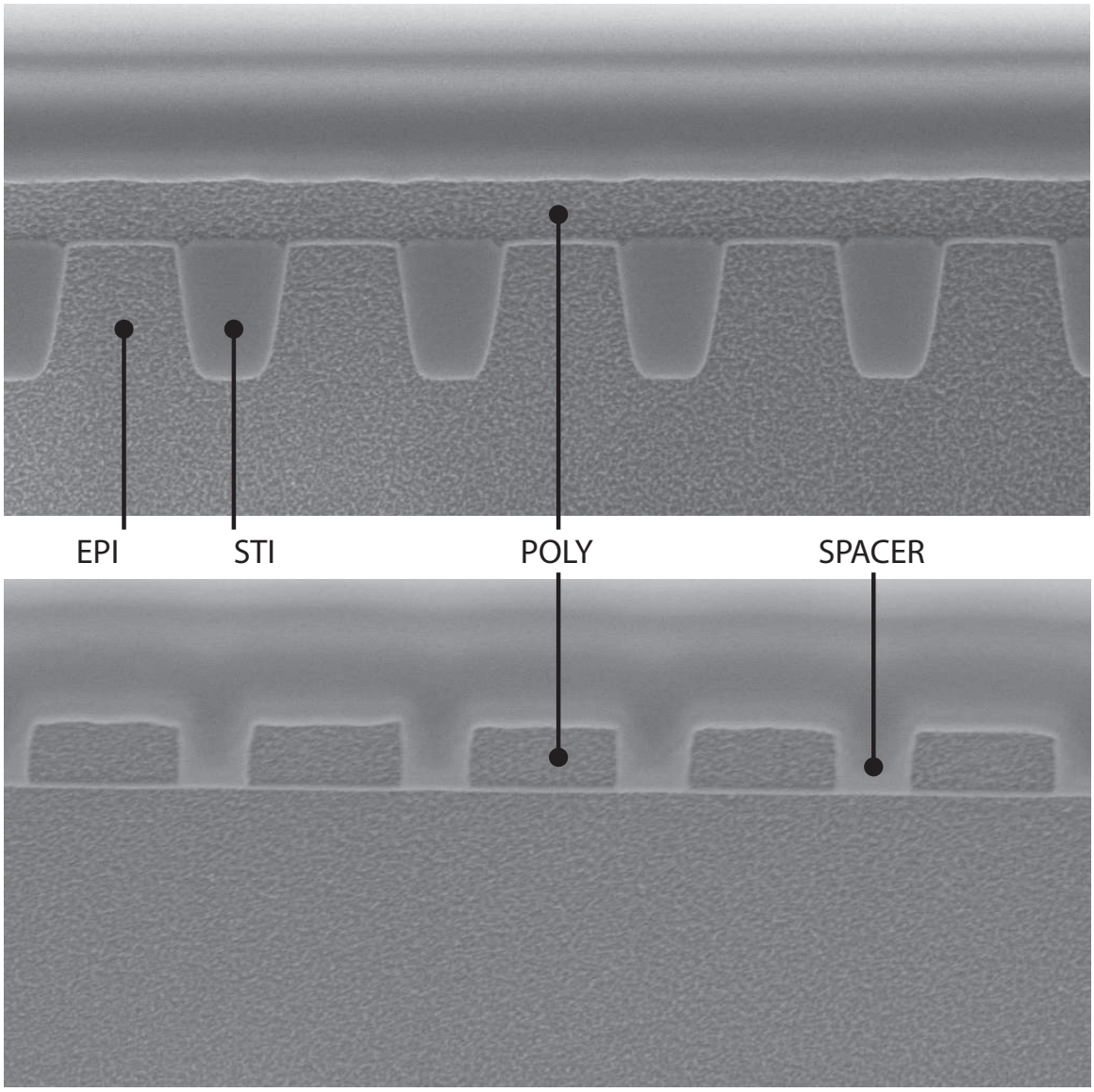



Figure 4.6: SEM cross-section across channel (top), and along channel (bottom).

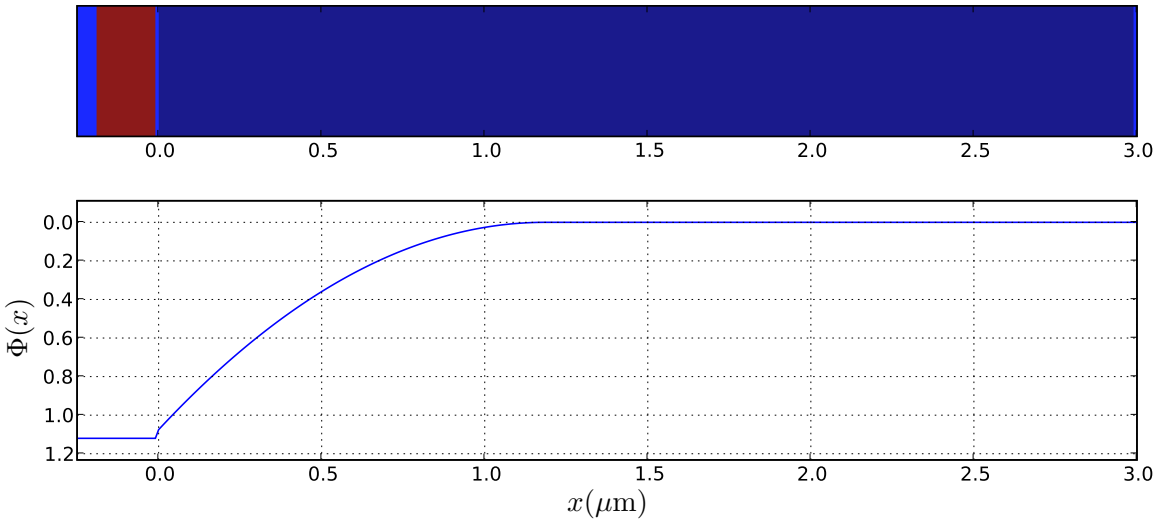



Figure 4.7: Surface-channel potential diagram.

plot in Figure 4.7. An n+ doped gate electrode is used with a bulk doping concentration of $N_A = 10^{15}/\mathrm{cm}^3$. Here the gate bias is held at 250mV which results in a potential well of over 1V. To eliminate the potential well, it is necessary to bring the electrodes to approximately -1V. Instead, the work function of the gate may be shifted by the energy band voltage of $E_G = 1.1$V if a p+ doped gate is used. Both gate types are used in the fabricated devices. The surface-channel device can store up to $\mathrm{C}_{ox}(V_g - V_{th})/q$ electrons, where $V_{th}$ is the threshold voltage of the MOS structure. Using $\mathrm{C}_{ox} = 5\mathrm{fF}/\mu\mathrm{m}^2$, and $V_{th} = 0$, nearly 100,000 electrons can be stored in a $1\mu\mathrm{m}^2$ at gate bias of 3V. The downside to the surface-channel device is that charge transport occurs along the Si-$\mathrm{SiO}_2$ interface where there are traps that degrade performance. In addition, dark current generated from the surface states are readily collected in the channel.

### 4.1.2 Buried-channel CCD

The purpose in implementing a buried-channel device is to move electrons away from the Si-$\mathrm{SiO}_2$ interface. This is accomplished by implanting a shallow n-type region under the electrode [46] [47]. Assuming that this region can be fully depleted by removing the free electrons, fixed positive ions from the donor atoms will be left behind. This increases the potential to a maximum that occurs deeper in the substrate. Photogenerated electrons are attracted to this region and, as long as the charge packets are sufficiently small, electrons will not interact with the Si-$\mathrm{SiO}_2$ interface during transport. The dose and energy specification

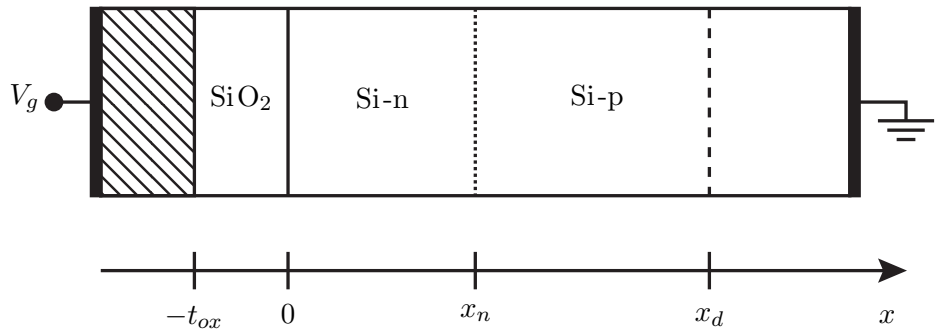



Figure 4.8: Buried-channel MOS structure.

for the implant is critical to the performance of the device. A one-dimensional analysis is given in the equations that follow. The structure for the buried-channel is shown in Figure 4.8. The electrode is polysilicon with a workfunction different from the bulk. The buried-channel implant extends to $x_n$ while the depletion region extends to $x_d$. From Gauss' law, we have

$$\frac{dE(x)}{dx} = \frac{\rho}{\epsilon} = \frac{qN(x)}{\epsilon_s} \tag{4.1}$$

$N(x)$ is the doping concentration for the semiconductor along $x$ moving into the bulk. In [7], $N(x)$ is used to for any 1-D doping concentration with donors represented as positive numbers and acceptors represented with negative numbers. We assume that free carriers are fully depleted along $N(x)$ up to $x_d$. This allows us to integrate the fixed charge in the depletion region to derive the equation for a buried channel.

$$E(x) = \int \frac{qN(x)}{\epsilon_s} dx \tag{4.2}$$

$$E(x) - E(x_d) = \int_{x_d}^{x} \frac{qN(x)}{\epsilon_s}; \qquad E(x_d) = 0 \tag{4.3}$$

$$E(x) = \frac{q}{\epsilon_s} \int_{x_d}^{x} N(x) dx \tag{4.4}$$

$$E(0) = \frac{q}{\epsilon_s} \int_{x_d}^{0} N(x) dx \tag{4.5}$$

In Equation 4.3, we apply the necessary boundary condition to find the electric field at the Si-$SiO_2$ interface. Since there is charge neutrality beyond $x_d$, the electric field is 0. We also know that $E(x)$ is constant across the gate dielectric ($SiO_2$). The electric field is greater in the gate dielectric than at the Si-$SiO_2$ interface as given by Equation 4.6. Substituting 4.5

into 4.6, gives the electric field in the gate dielectric region.

$$E_{ox} = \frac{\epsilon_s}{\epsilon_{ox}} E(0) \tag{4.6}$$

$$E_{ox} = \frac{q}{\epsilon_{ox}} \int_{x_d}^{0} N(x)dx \tag{4.7}$$

Next, we find the surface potential $\Phi_s$ at the Si-$SiO_2$ interface by defining the boundary conditions and integrating over the range of $-t_{ox} < x < 0$. For a given doping profile, $\Phi_s$ is determined by the voltage applied to the electrode $V_g$ and the depletion width $x_d$ as found in Equation 4.12.

$$\frac{-d\Phi(x)}{dx} = E_{ox}; \qquad -t_{ox} < x < 0 \tag{4.8}$$

$$\Phi(x) = -\int E_{ox}dx; \qquad -t_{ox} < x < 0 \tag{4.9}$$

$$\Phi(-t_{ox}) - \Phi(0) = -\int_{0}^{-t_{ox}} E_{ox}dx \tag{4.10}$$

$$V_g - \Phi_s = E_{ox}t_{ox} \tag{4.11}$$

$$\Phi_s = V_g + \frac{q}{\epsilon_{ox}} \int_{0}^{x_d} N(x)dx \tag{4.12}$$

Now we use Equation 4.4 to solve for $\Phi(x)$, again applying the boundary conditions and this time using integration by parts.

$$\Phi(x) = -\int E(x)dx; \qquad x > 0 \tag{4.13}$$

$$\Phi(x) - \Phi(0) = -\int_{0}^{x} E(x)dx \tag{4.14}$$

$$\Phi(x) - \Phi(0) = -xE(x) + \int_{0}^{x} xdE \tag{4.15}$$

$$\Phi(x) - \Phi(0) = \frac{qx}{\epsilon_s} \int_{x}^{x_d} N(x)dx + \frac{q}{\epsilon_s} \int_{0}^{x} xN(x)dx \tag{4.16}$$

$$\Phi(x) = \Phi_s + \frac{qx}{\epsilon_s} \int_{x}^{x_d} N(x)dx + \frac{q}{\epsilon_s} \int_{0}^{x} xN(x)dx \tag{4.17}$$

$$\Phi(x) = V_g + \frac{q}{\epsilon_{ox}} \int_{0}^{x_d} N(x)dx + \frac{q}{\epsilon_s} \int_{0}^{x} xN(x)dx + \frac{qx}{\epsilon_s} \int_{x}^{x_d} N(x)dx \tag{4.18}$$

The real value for $V_g$ depends on the different doping levels of the polysilicon electrode and the bulk substrate due to the resulting work function difference. This is defined by:

$$\Phi_{MS} = \Phi_M - \Phi_{Si} \tag{4.19}$$

where $\Phi_M$ is the work function of the electrode and $\Phi_{Si}$ is the work function of the bulk silicon. Since both the gate electrode and the bulk are made of silicon, we can write the work function difference in terms of Fermi levels.

$$\Phi_{MS} = (E_i - E_F)_{gate} - (E_i - E_F)_{bulk} \tag{4.20}$$

$$(E_i - E_F) = \frac{KT}{q} ln\left(\frac{N_A}{n_i}\right) \qquad \text{for p-type dopant} \tag{4.21}$$

$$(E_i - E_F) = -\frac{KT}{q} ln\left(\frac{N_D}{n_i}\right) \qquad \text{for n-type dopant} \tag{4.22}$$

Usually the gate electrode is degenerately doped n-type or p-type and the bulk is p-type. In this case, the levels can be written as follows.

$$(E_i - E_F) = \frac{E_G}{2} \qquad \text{for p+ dopant} \tag{4.23}$$

$$(E_i - E_F) = -\frac{E_G}{2} \qquad \text{for n+ dopant} \tag{4.24}$$

With a bulk doping level of $N_A = 10^{15}/\text{cm}^3$, $\Phi_{MS}$ is 0.25V for a p+ electrode and -0.85V for an n+ electrode. The final form for the potential of the buried-channel equation is given in Equation 4.25.

$$\Phi(x) = V_g - \frac{KT}{q} ln\left(\frac{N_A}{n_i}\right) \pm \frac{E_G}{2} + \frac{q}{\epsilon_{ox}} \int_0^{x_d} N(x)dx + \frac{q}{\epsilon_s} \int_0^{x} xN(x)dx + \frac{qx}{\epsilon_s} \int_x^{x_d} N(x)dx \tag{4.25}$$

We can now use the equation to determine the initial design for the buried-channel device. We find that it is necessary to use a p+ electrode so that large negative voltages are not required. With a zero bias on $V_g$, we set the channel potential to be approximately 1V so that a floating diffusion readout circuit can induce full depletion of the channel. We approximate the channel with a constant doping profile of $x = 300\text{nm}$ and $N_D = 2 \times 10^{16}/\text{cm}^3$. A plot of Equation 4.25 with these parameters is shown in Figure 4.9.

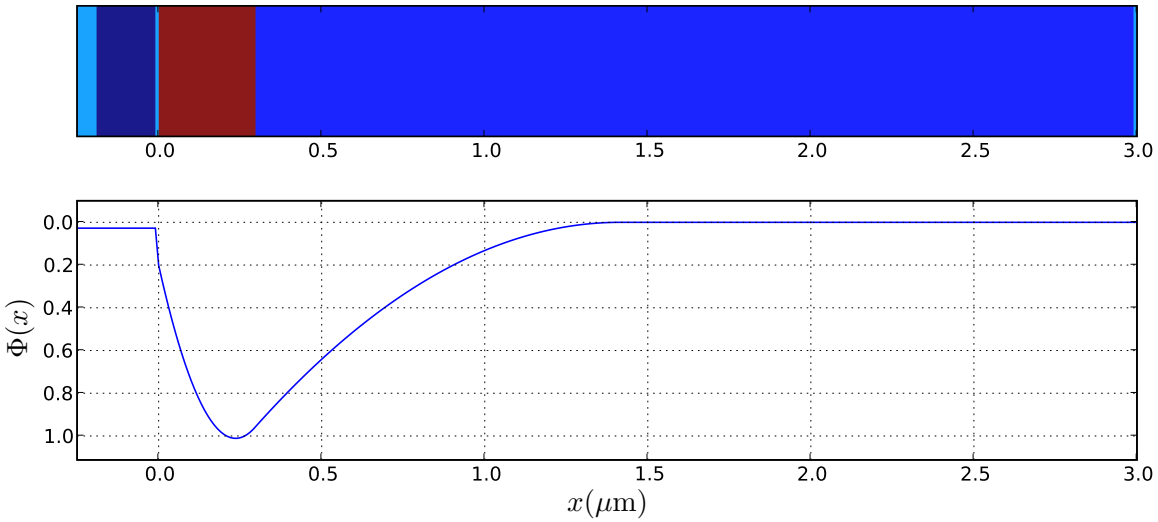



Figure 4.9: Buried-channel potential diagram.

### 4.1.3 Pinned phase buried-channel CCD

The buried-channel device improves upon the surface-channel device by moving the charge transport away from the Si-$SiO_2$ interface. However, the buried-channel device offers little to no benefit over the surface-channel device in terms of dark current performance since generation of carriers from the surface states are still attracted into the buried channel. The pinned phase buried-channel device uses an extra implant between each electrode that allows the surface to be inverted during integration and yet still contain the barrier necessary for charge confinement. The inverted channel contains a large concentration of holes which make an efficient recombination region for the surface generated carriers and acts as a shield against dark current. The pinned phase buried-channel design is shown in Figure 4.10 with doping profiles plotted in Figure 4.11-4.12. This design is similar to the open-pinned phase CCD described in [48] where the channel is inverted during the integration time. Instead of integrating charge under the P-type implant, we integrate charge under the electrode with an inverted surface. During the integration time, the entire surface is pinned with a large concentration of holes provided by the channel stops, which reduces the dark current at the interface.

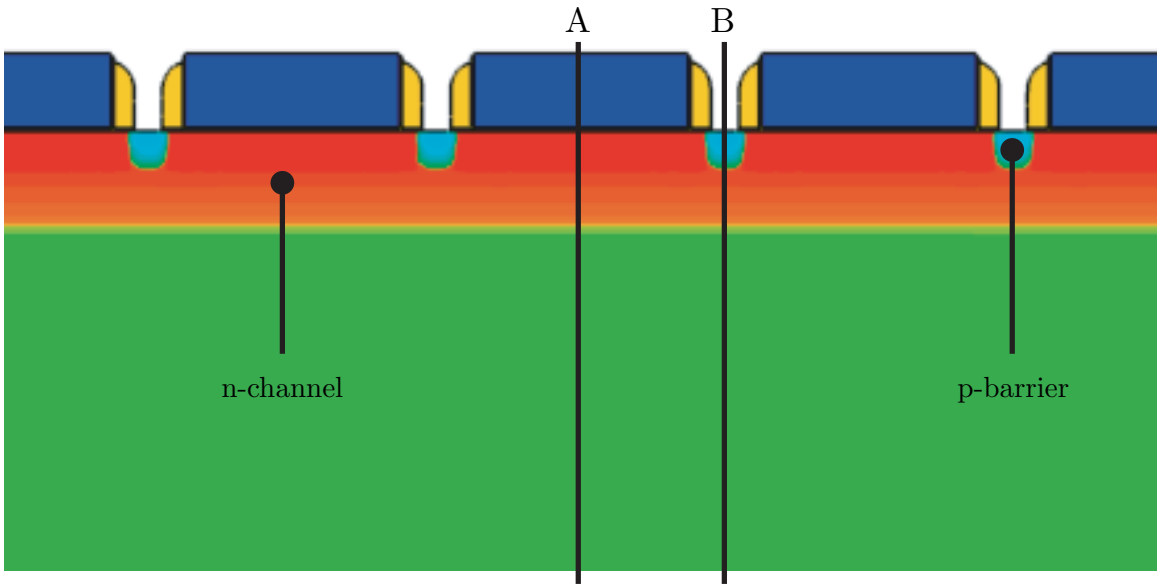



Figure 4.10: Design of the pinned phase buried-channel CCD showing P+ electrodes with n-type channel along section A and self-aligned P-type barriers along section B.

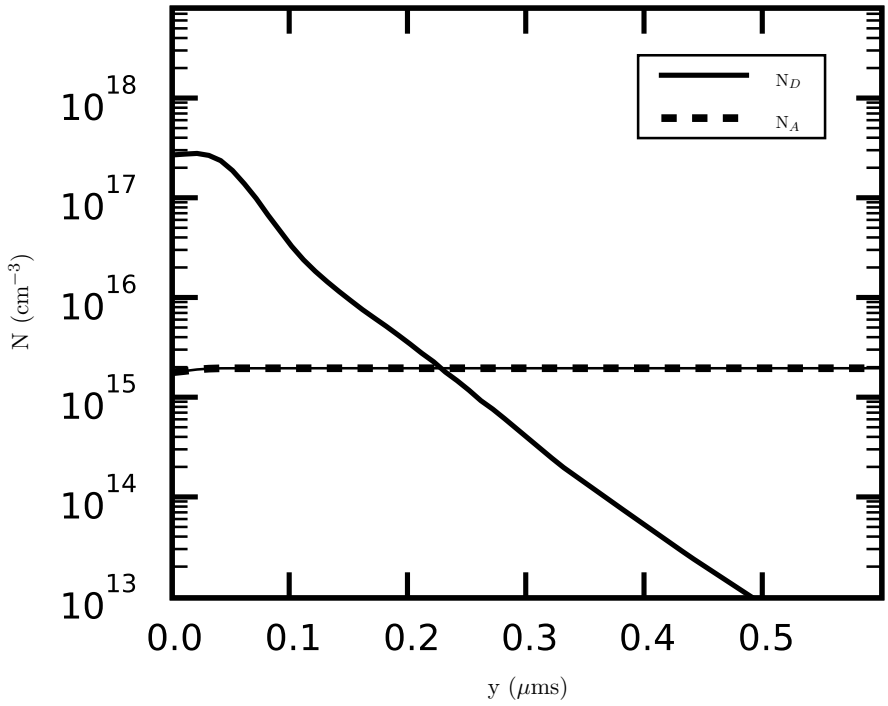



Figure 4.11: Simulated pinned phase doping profile along cross-section A.

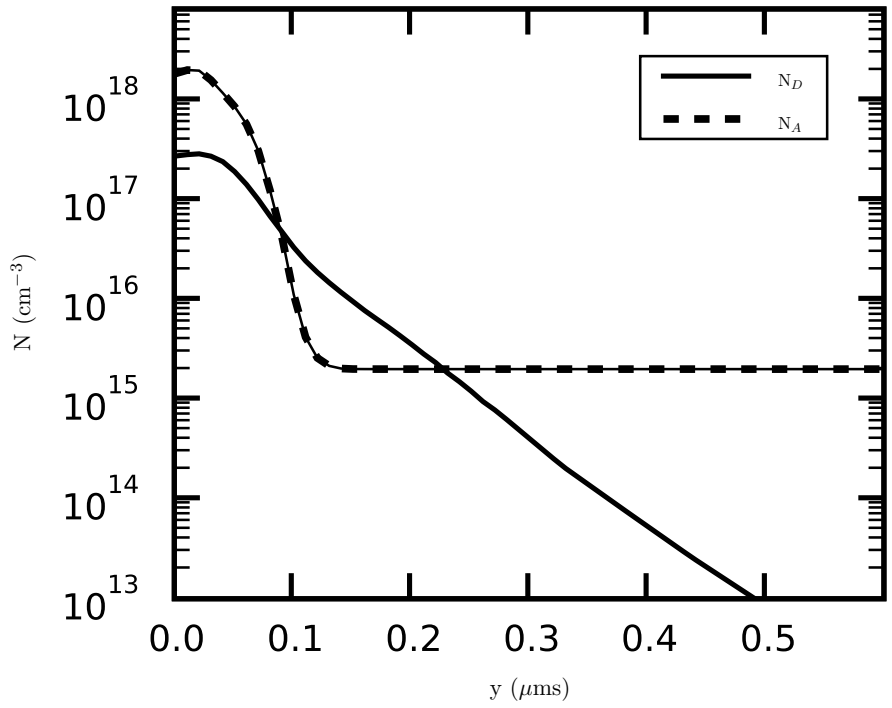


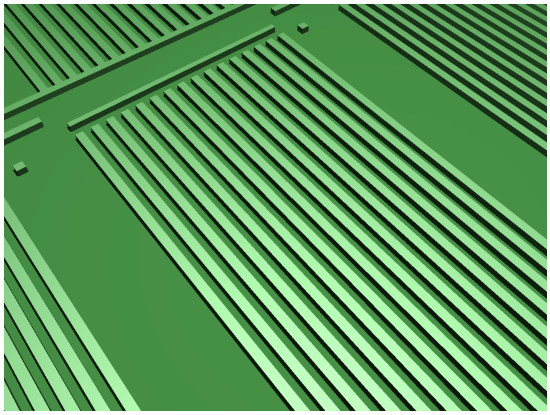

Figure 4.12: Simulated pinned phase doping profile along cross-section B.

## 4.2 Fabrication Process

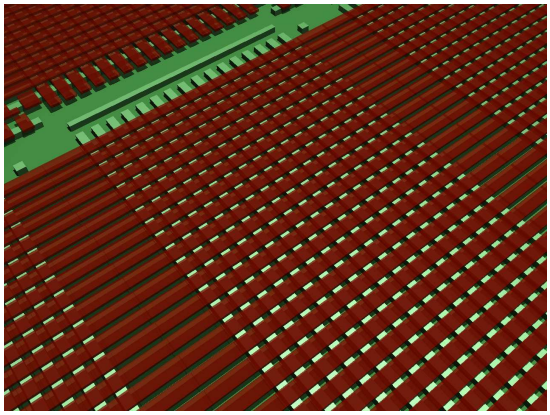
Each pixel consists of a single poly electrode, a channel, and a channel stop. The channels and stops for the surface-channel devices are shown in Figure 4.13a. The electrodes are patterned with non-silicided polysilicon as shown in Figure 4.13b. Control lines are wired with the first metal layer as shown in Figure 4.13d. Each pixel array is separated by a wall of 4 metal layers. Processing up to the second metal layer is shown in Figure 4.13f.

Additional steps were added to the CMOS process to create the CCD designs. Figure 4.14 shows a simplified version of the process flow that was developed. The plots are generated from Sentaurus TCAD process simulations to show how the horizontal CCD is implemented. We used the simulator to define the implant recipe for the buried channel, the channel stops and the self-aligned pinned-phase implant. We use a P-type epi substrate with an acceptor concentration of $10^{15}/\text{cm}^3$ and thickness of 3.0um. The bulk wafer substrate has an acceptor concentration of $10^{19}/\text{cm}^3$. We start in Figure 4.14a by thermally growing a thin layer of sacrificial screening oxide for about an hour in flows of oxygen and nitrogen. Next is the Pwell implantation shown in Figure 4.14b. We anneal the damage during implantation and drive the dopants deeper into the substrate. Figure 4.14c shows the buried channel implant. Several splits were used to determine the best concentration for this layer as described later. Figure 4.14d shows the polysilicon deposition and etch followed by the reoxidation in Figure 4.14e. Before spacers are placed, the S/D extension regions for the readout transistor are defined in Figure 4.14f. Nitrided spacers are placed in Figure 4.14g. The P and N-type S/D implants are shown in Figure 4.14h and Figure 4.14i. An extra step is used to block silicide from forming on the active electrodes as shown in Figure 4.14j and Figure 4.14k.

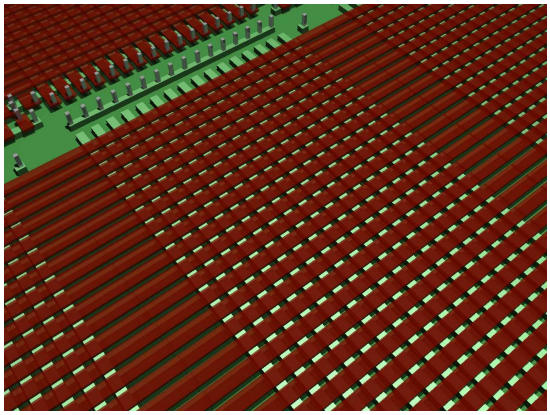
In all designs, the polysilicon is doped by masking out the channels as shown in Figure 4.15. The resolution of the S/D implant masks were the limiting factor in scaling beyond 0.5$\mu$m pixel pitch at this process node. We expect that it is possible to pre-dope the polysilicon before etch to scale the pixel size further. The polysilicon for the surface-channel device is doped N+ and the buried-channel designs are doped P+ to shift the workfunction closer to the operating range of CMOS circuits. The IOs on the test chip were designed to allow both positive and negative voltage sequencing. An electrode spacing of 180nm was used in all designs. The polysilicon is 130nm thick with gate oxide of 8nm. The channel stop for the surface-channel device is Shallow Trench Isolation (STI). The channel stop for the buried-channel is formed by a P-type implant (BF2, 75keV, 4.0E13/cm$^2$). The SEM



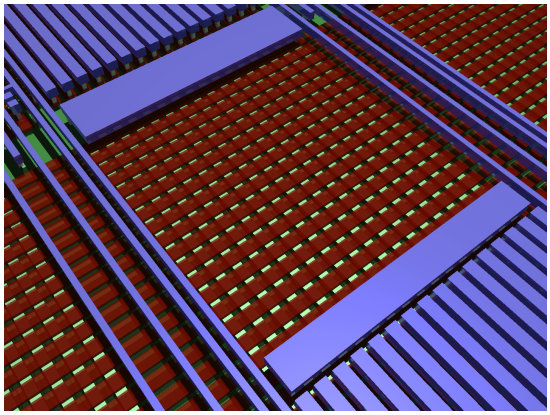
(a) Channel stops.



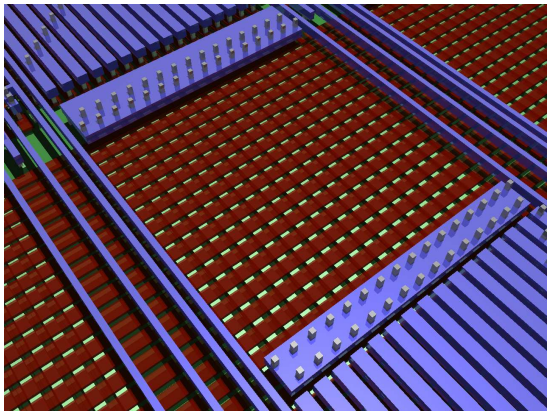
(b) Polysilicon.



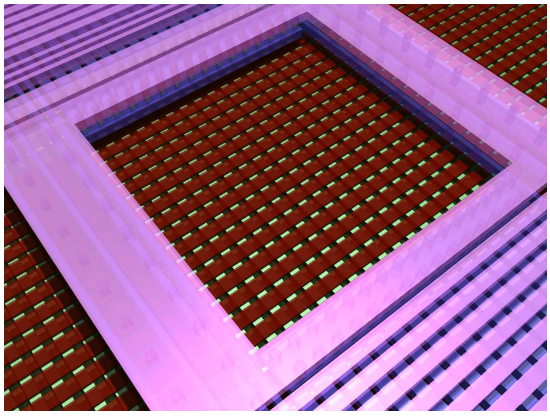
(c) Contacts.



(d) Metal1.

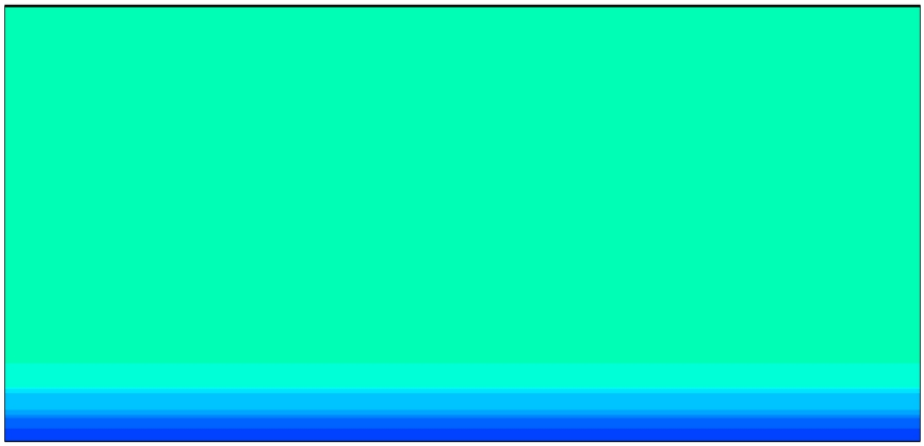


(e) Via for Metal1/Metal2.

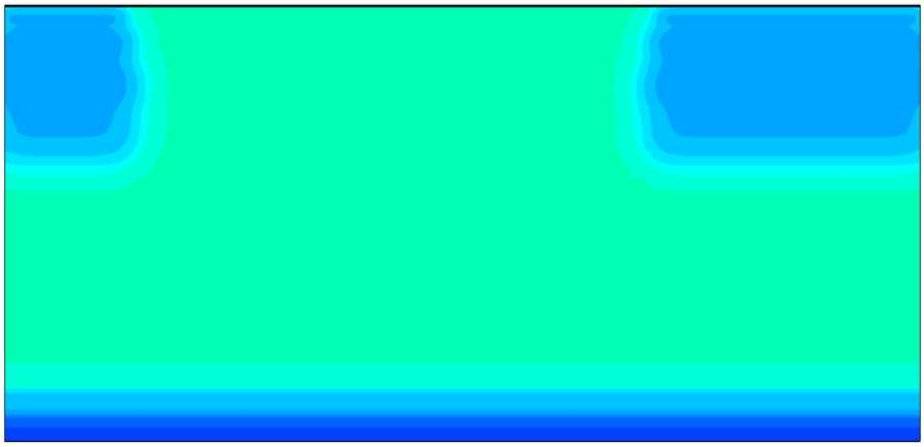


(f) Metal2.

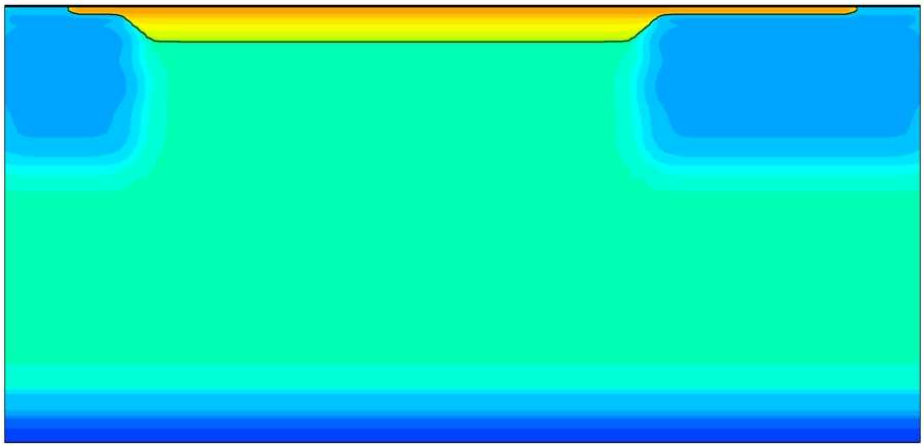
Figure 4.13: Structure of FT-CCD.



(a) Epi substrate with screening oxide.

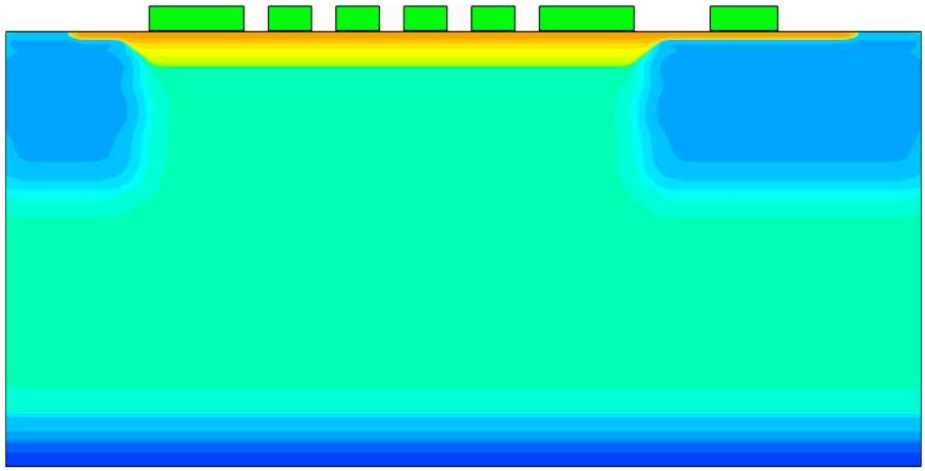


(b) Pwell implantation and activation.

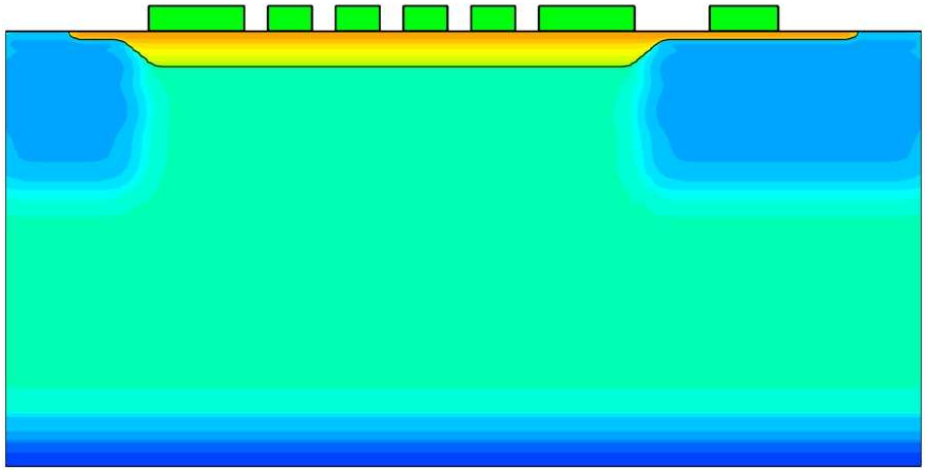


(c) N-type buried implantation.

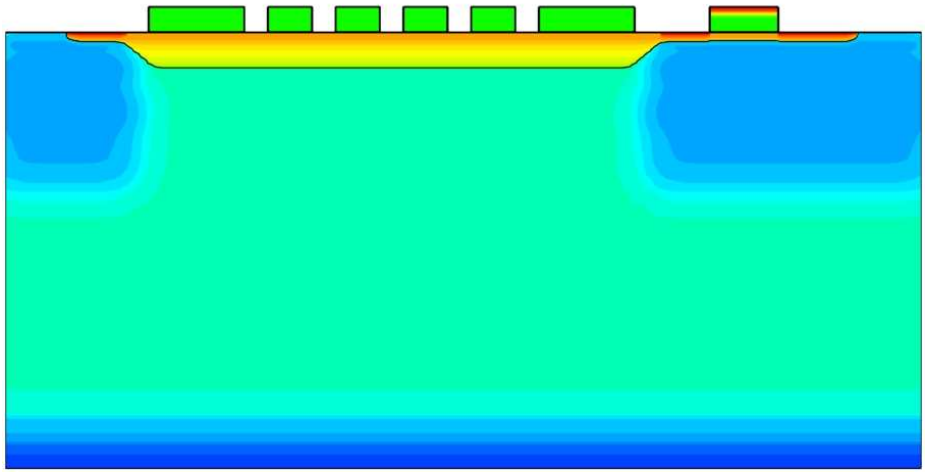
Figure 4.14: Buried channel CCD process flow.



(d) Polysilicon deposition and etch.

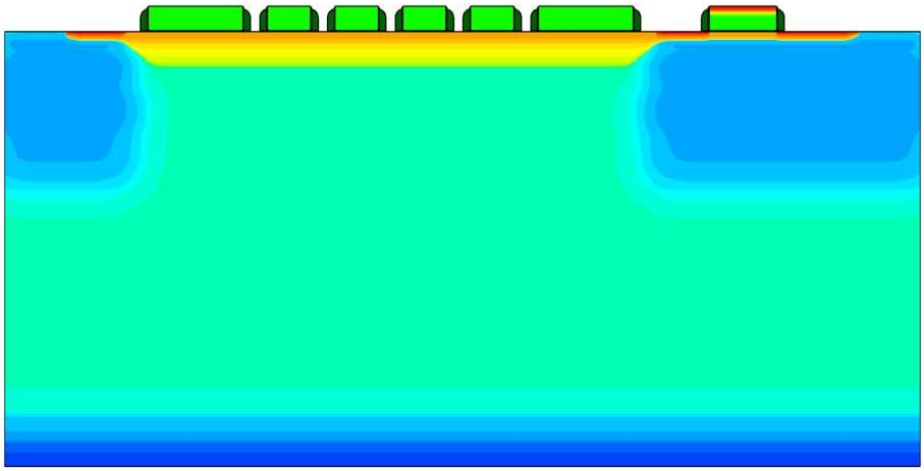


(e) Polysilicon reoxidation.

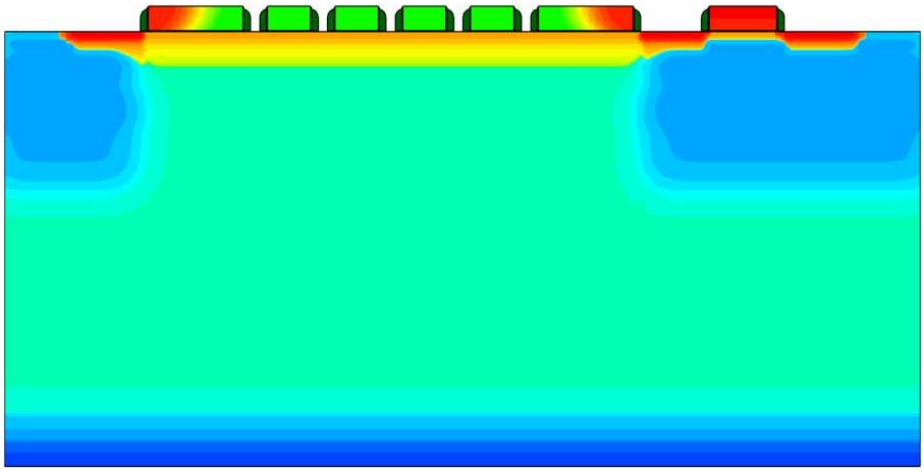


(f) S/D extension implantation.

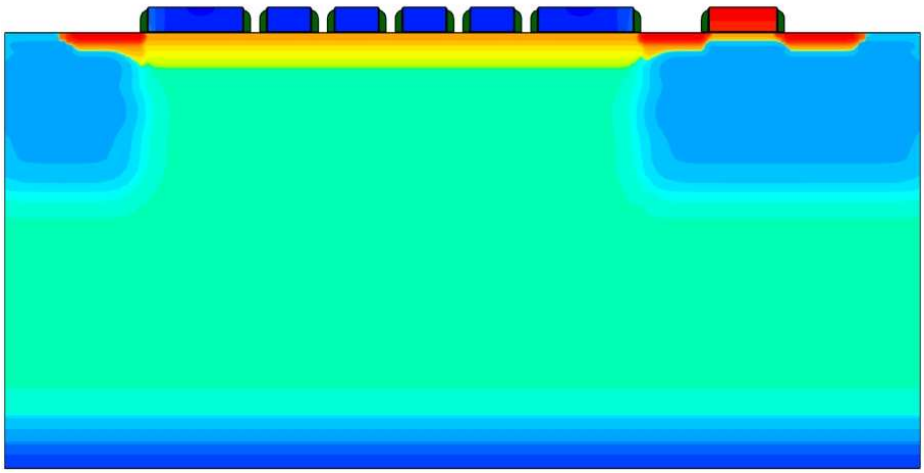
Figure 4.14: Buried channel CCD process flow. (cont.)



(g) Nitride spacers.

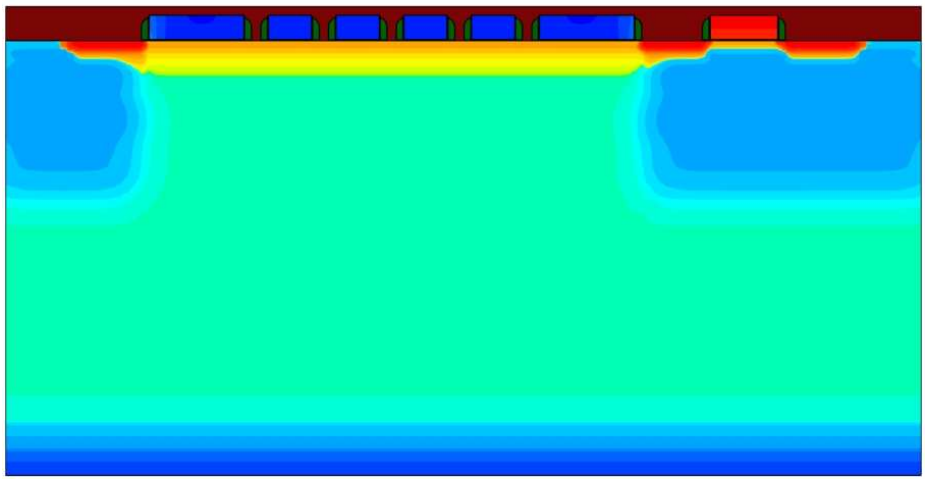


(h) N-type S/D implantation.

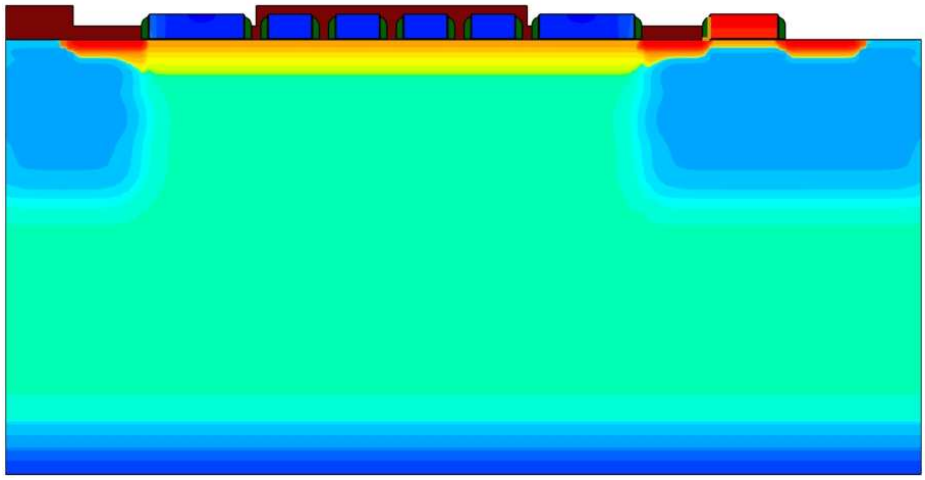


(i) P-type S/D implantation.

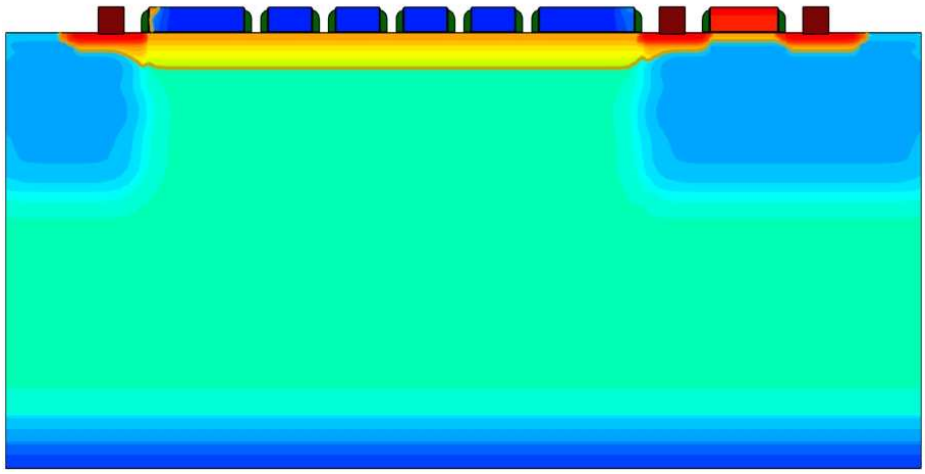
Figure 4.14: Buried channel CCD process flow. (cont.)



(j) Photoresist for silicide block.



(k) Selective photoresist for silicided polysilicon.



(l) Tungsten contacts.

Figure 4.14: Buried channel CCD process flow. (cont.)

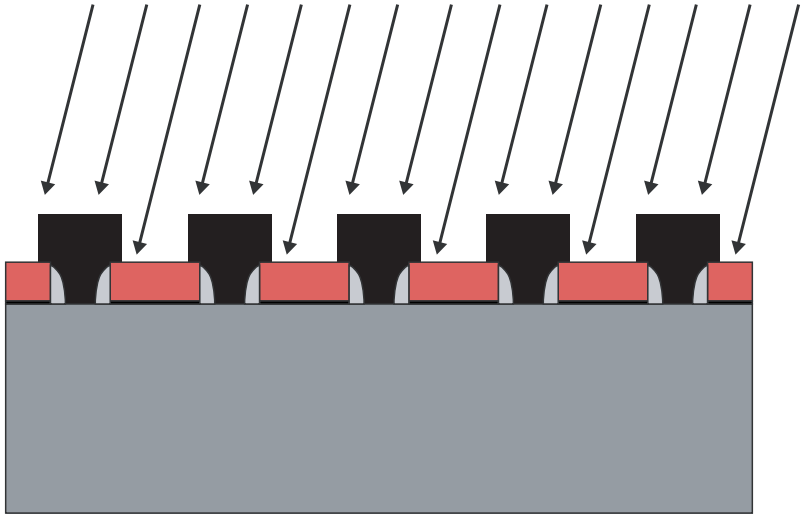


Figure 4.15: Method for doping the polysilicon such that the channel region is protected.

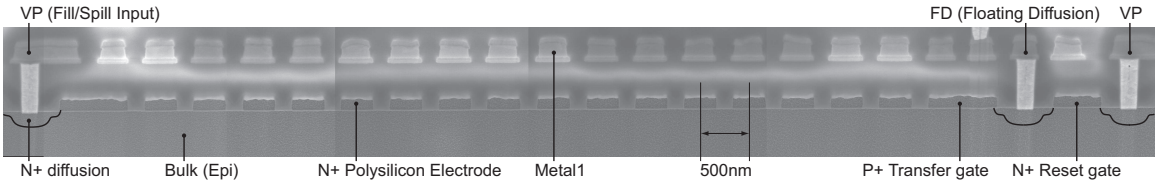



Figure 4.16: SEM of 16-stage H-CCD showing fill/spill input for electrical testing, floating diffusion for output charge-to-voltage conversion, and the reset gate.

for the H-CCD with fill-and-spill input and floating diffusion is shown in Figure 4.16.

An image is captured by integrating photocharge at each electrode or at every other electrode for higher well capacity. The integration begins by depleting the CCDs of charge via transfer to the upper diffusion V0. During integration, the pixel array electrodes are held at an intermediate voltage. At the end of integration, the accumulated charge is ripple transferred row-by-row to the storage array and then into the H-CCD one pixel at a time until every pixel has been double sampled at the floating diffusion and buffered by the source follower transistor.

## 4.3 Simulation and Measurement Results

Simulated potential diagrams along the channel (wherever the max potential occurs) for several phases are shown for all 3 designs in Figure 4.17-4.19. Single electrode charge confinement is achieved in the surface-channel device due to the barriers created by the

poly gap spacing, whereas it is achieved in the buried-channel device due to the induced pockets. The pinned phase design uses the self-aligned P-type implants as barriers to confine the charge. Although the surface potential is pinned to the channel stop potential, the depleted channel under the electrode remains at a higher potential. With sufficient gate voltage, each of the designs overcome the pocket or barrier that creates the confinement and charge is transferred away from one region and then packed again at single electrode pitch into another.

Confining charge at every electrode is made possible by using ripple charge transfer. Since we build distributed subarrays at the stack height scale, the overhead of employing ripple charge transfer is minimal. One drawback of the single electrode charge confinement is the reduced well capacity of about 500e-. We show that we can boost the well capacity by an order of magnitude by running in an interlaced mode where every other electrode induces a large barrier to confine more charge. By patterning the polysilicon at half the channel stop pitch, we could implement this method of operation at our current pixel scale without interleaving. At the current process node, this would require pre-doping the polysilicon.

Charge transfer efficiency is measured highest for the surface-channel device at just above 99.9%, whereas the buried-channel device has to be carefully tuned for this level of performance. When the devices show comparable CTE performance, the surface-channel also shows a slightly lower dark current than the buried-channel device. We also fabricated the surface-channel device with P-type channel stops and found that the CTE degrades. When the surface channel electrodes are first accumulated with holes, the CTE drops to as low as 98%, whereas the buried-channel device does not degrade under this condition. Although we can measure the CTE just after attracting the holes to the surface, we cannot operate the buried-channel device in this way during the integration period because the charge confinement region is eliminated as the surface is pinned. Therefore, the buried-channel device offers little to no benefit over the surface-channel device in terms of dark current performance. For this reason, we implemented the pinned phase buried-channel device that uses an extra implant between each electrode. This allows the surface to be inverted during integration and yet still contain the barrier necessary for charge confinement. The barrier potential relative to the channel potential under the electrode is shown in Figure 4.20. The applied electrode voltages for V2 and V3 are labeled for each case. During the integration cycle, we need the potential under the electrode to be higher than the barrier potential in order to confine the electrons at the electrode. We also require that the surface

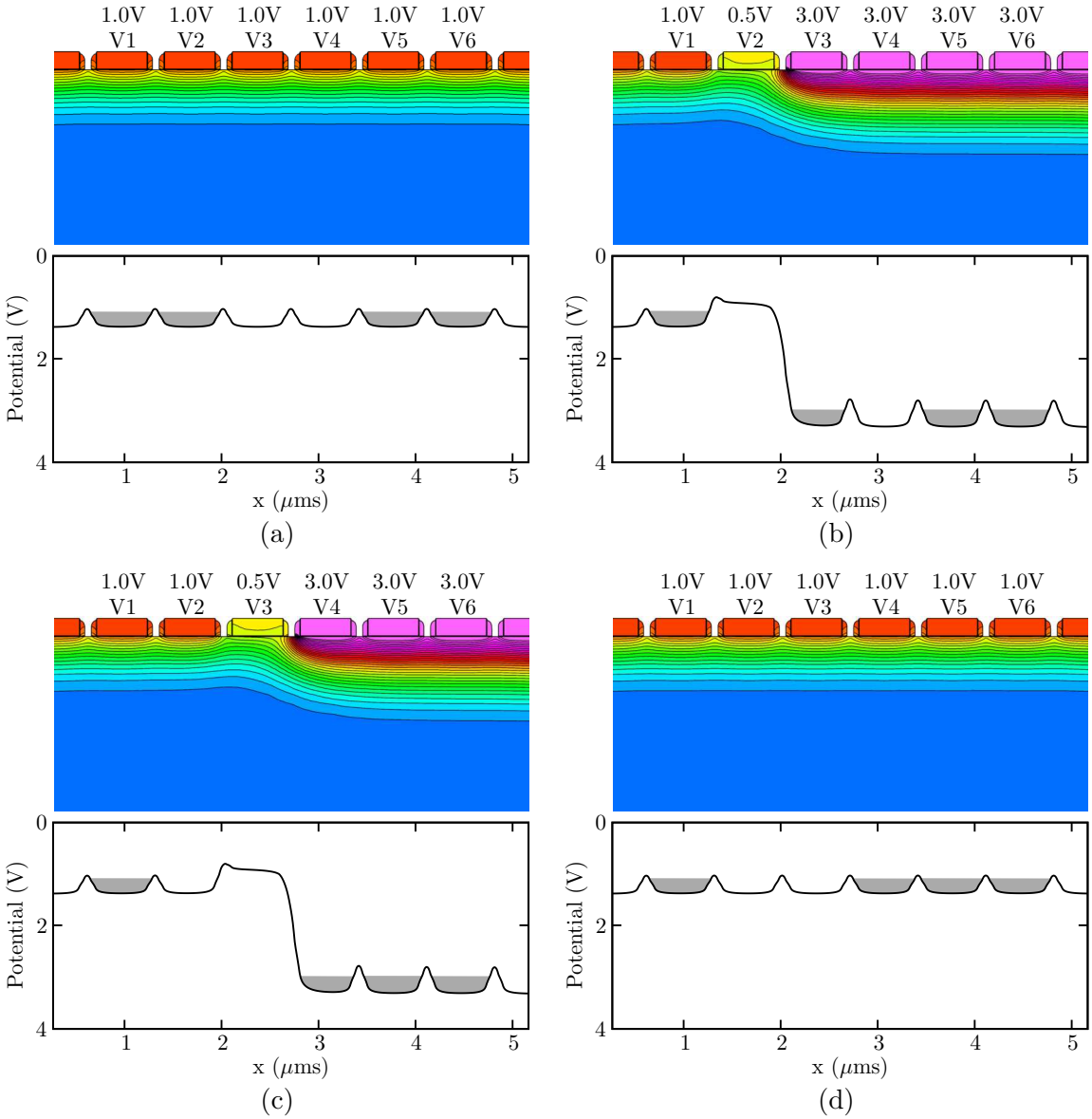



Figure 4.17: Potential diagrams for the surface-channel CCDs with single electrode charge confinement during ripple charge transfer.

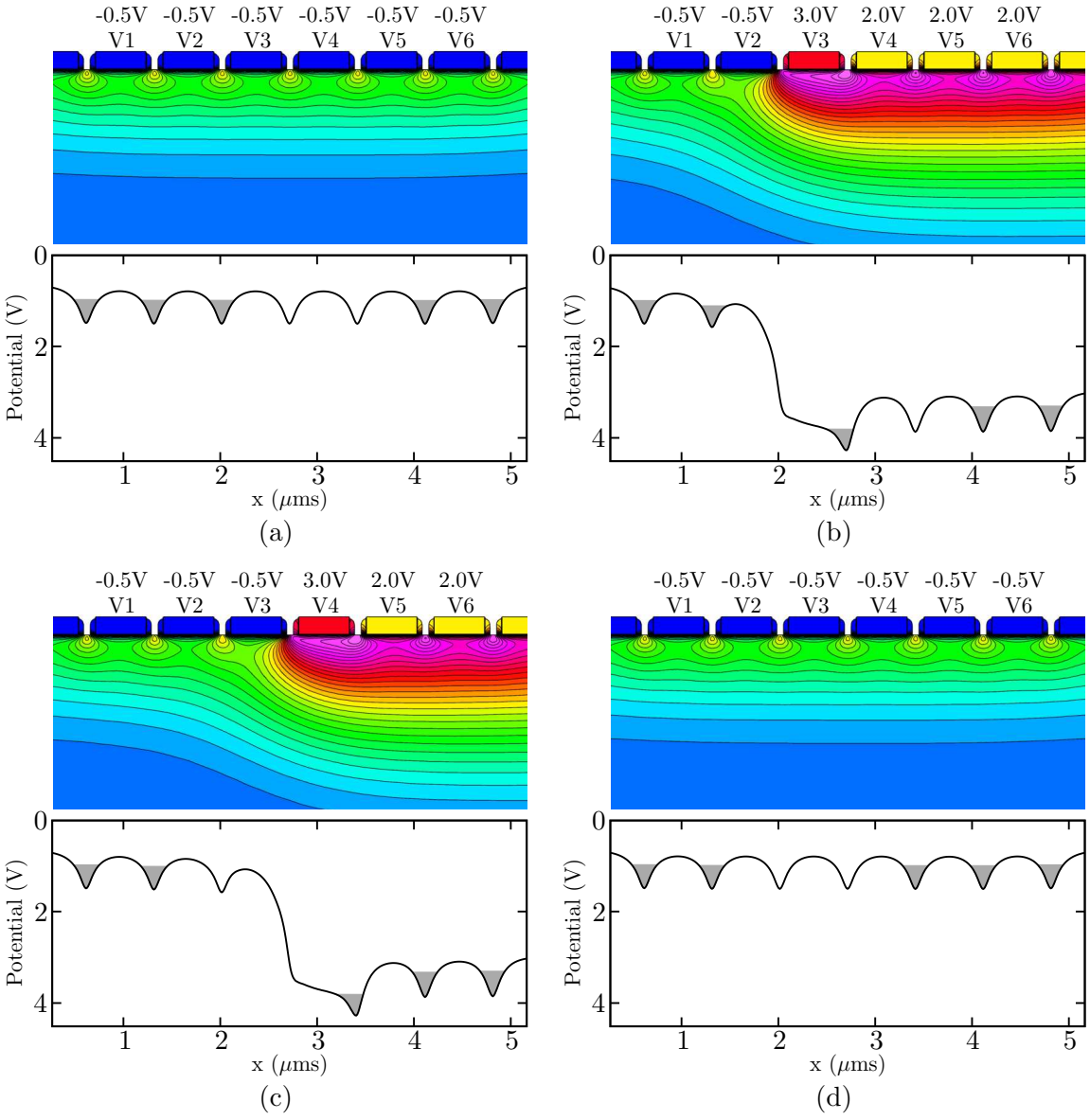



Figure 4.18: Potential diagrams for the buried-channel CCDs with single electrode charge confinement during ripple charge transfer.

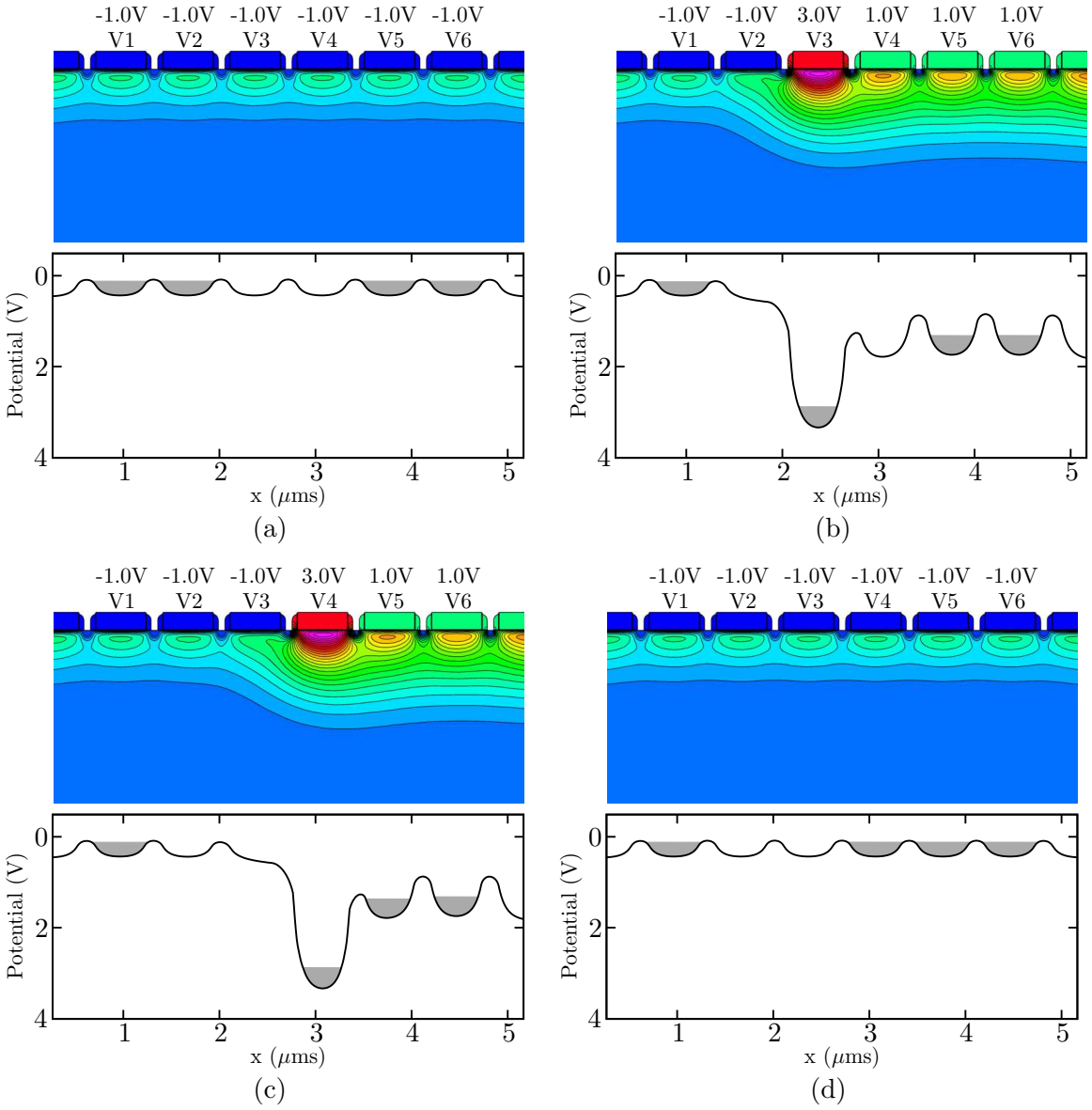



Figure 4.19: Potential diagrams for the pinned phase buried-channel CCDs with single electrode charge confinement during ripple charge transfer.

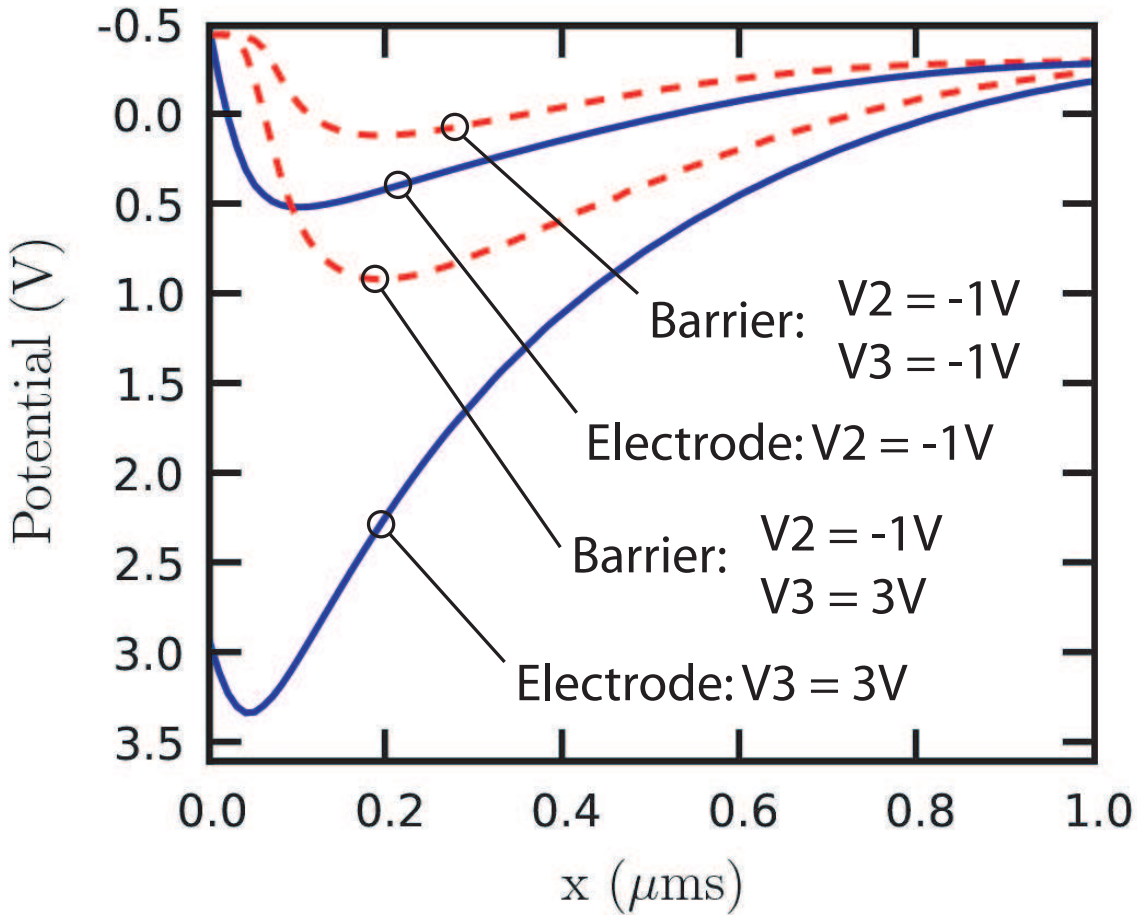



Figure 4.20: Pinned phase diagram showing the barrier potentials relative to the potentials under the electrode region.

potential be low enough to substantially increase the hole concentration. Since the inter-electrode implants are of P-type dopant, their potentials are lower than the buried channel under the electrodes even when the electrodes are low enough to pin the surface. This is the case when all electrodes are held at -1V. We are able to manipulate the barrier potential by increasing the potential of the adjacent electrode. When V3 is set to 3V, the barrier at V2 becomes substantially lower than the potential under V2, which causes the charge to transfer from one electrode to the next.

The conversion gain for the 0.5 $\mu$m pixel is 193$\mu$V/e- and 165$\mu$V/e- for the 0.7 $\mu$m pixel. There is no significant gain difference between the 3 types of CCDs because the readout transistors are identical. Despite the use of poly electrodes, the quantum efficiency (QE) is reasonable for short wavelengths as shown in Figure 4.21. This is due to the thin the poly layer and the open space in between each electrode. The dark current is about 35e-/sec for both the surface and buried-channel devices The dark current improves by a factor of 15 for the pinned phase device. When operating in the pinned phase mode, the well capacity cannot be adjusted once the surface is pinned. We used a 4-way split on the n-channel doping profile in order to obtain a reasonable range of performance. Only the lower dopant levels were successful, which may demonstrate some limitations in the achievable well capacity with this approach.

A photomicrograph of the test chip is shown in Figure 4.22 with a magnified view of the aperture array containing 0.5$\mu$m pixels. A test board was designed with an FPGA and an array of Digital-to-Analog Converters (DACs) to provide all the sequencing to the CCD chip. The chip was bonded to a ceramic PGA package and inserted into a ZIF socket for testing as shown in Figure 4.23.

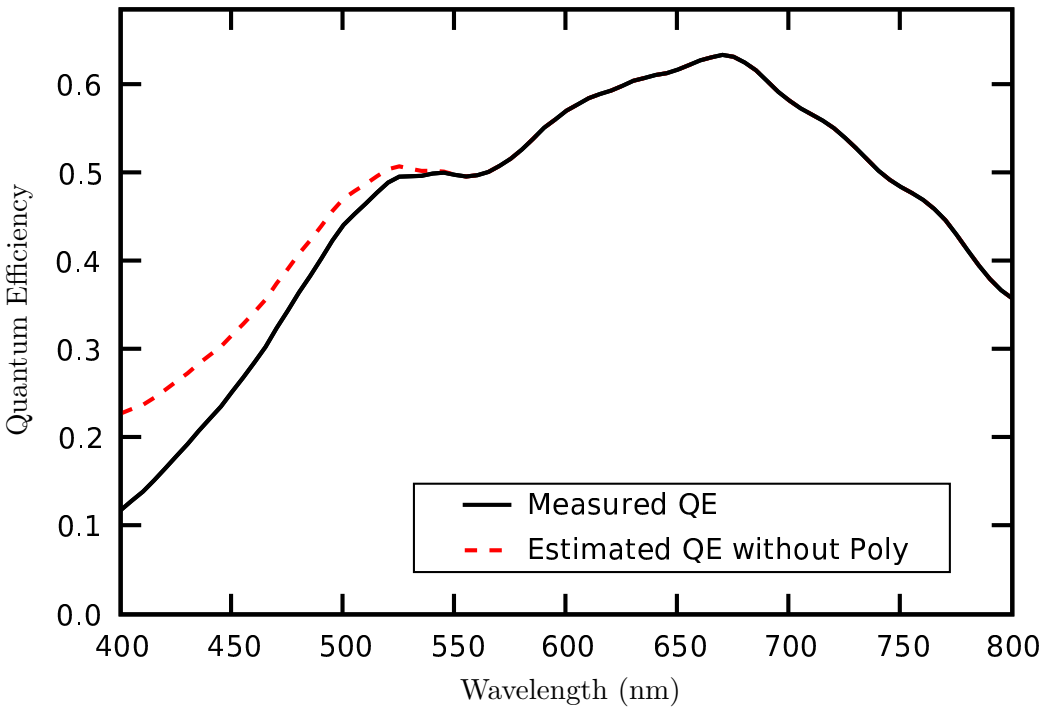
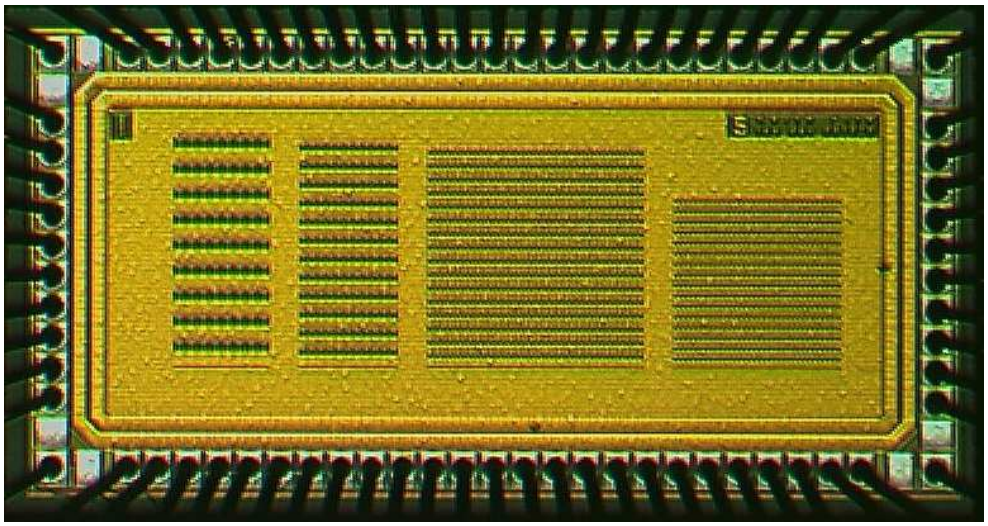
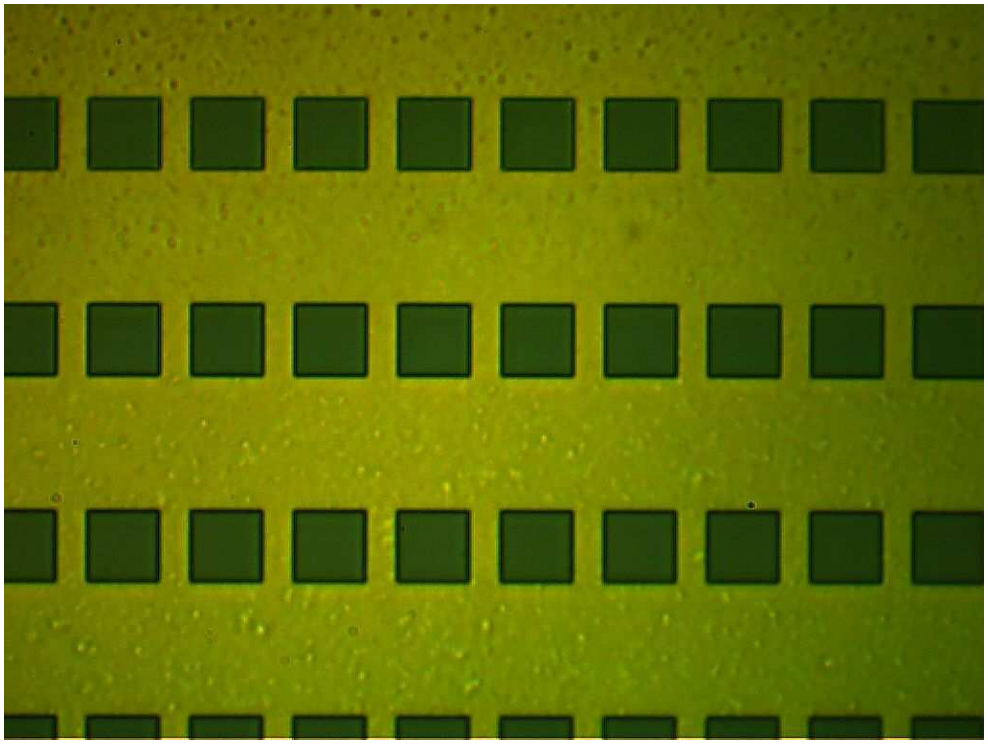



Figure 4.21: Measured quantum efficiency for the buried-channel CCD.

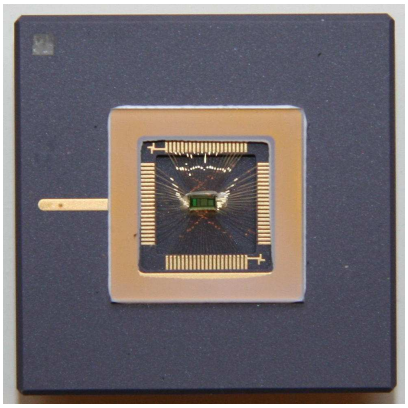


(a)

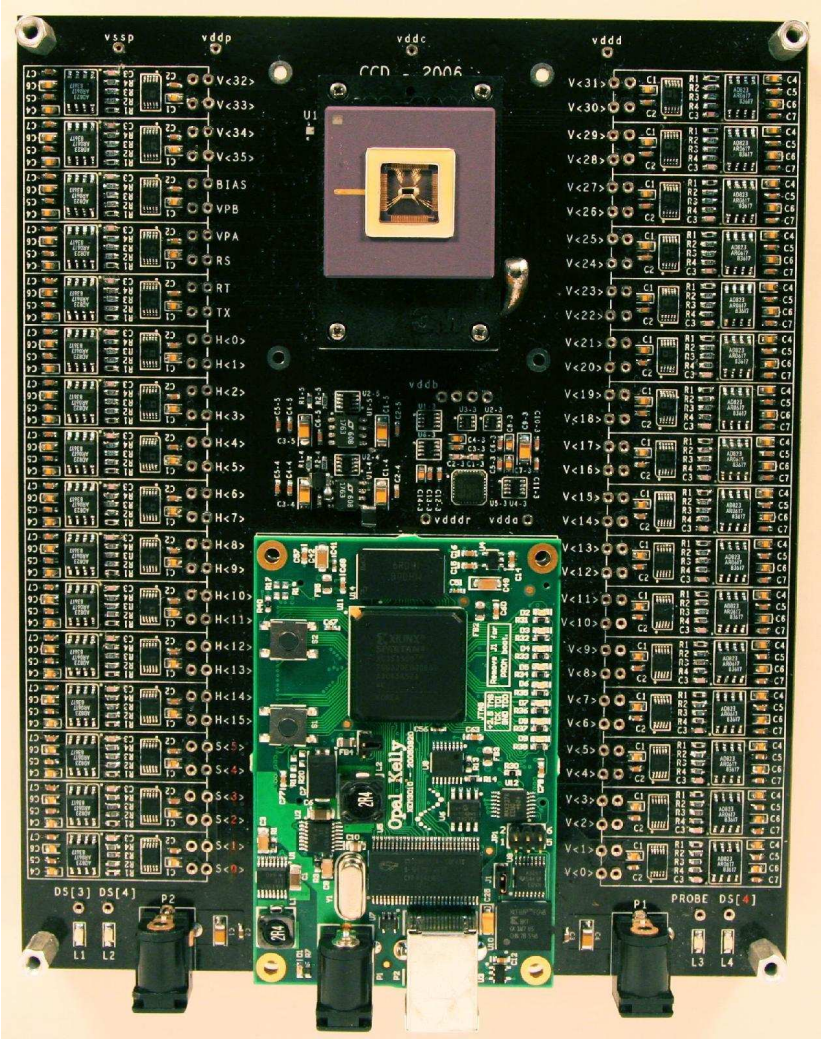


(b)

Figure 4.22: (a) Test chip micrograph and (b) magnified view of aperture array.



(a)




(b)

Figure 4.23: (a) Packaged test chip and (b) test board for CCD sequencing and data capture.

# Chapter 5

# Multi-Aperture Sensor Design

A block diagram of the multi-aperture image sensor is shown in Figure 5.1. The chip comprises a 166×76 aperture array, each with a 16×16 FT-CCD. The aperture control buses, V[35:0] and H[15:0] are globally connected to the FT-CCD array. To select a row of FT-CCD readout circuits, the RS signal is applied through the row decoder addressed by the ROW signal. The MUX blocks contain column control circuits, bias circuits, inputs for external testing of each column analog chain, and support for serial analog pixel readout through AOUT. The per-column ADCs share an output bus, which is controlled by the signal COL and buffered for digital readout through DOUT[10:0]. After every conversion cycle, the digitized values are read out out from the ADC buffers one column at a time. An ESD clamp to the lowest chip potential, typically used on IO pins, is not implemented on the CCD control lines so that negative voltages can be applied during testing and characterization.

In the following section, we describe the design, fabrication and operation of the FT-CCD. In Section 5.2, we describe the chip operation, and in Section 5.3, we describe the circuit and operation of the per-column ADC.

## 5.1 Aperture Array

The schematic and device cross-sections of the 16×16 pixel FT-CCD are shown in Figure 5.2. It consists of a pixel array, a frame buffer, a horizontal CCD (H-CCD) with floating diffusion (FD), and follower readout. The CCDs are formed using p+ polysilicon electrodes, n-type CCD channel and p-type channel stop implants. A total of 4 customized implants are needed to implement this structure in a standard process. The purpose of the additional implants is to improve charge transfer efficiency by using a buried channel instead of a surface channel

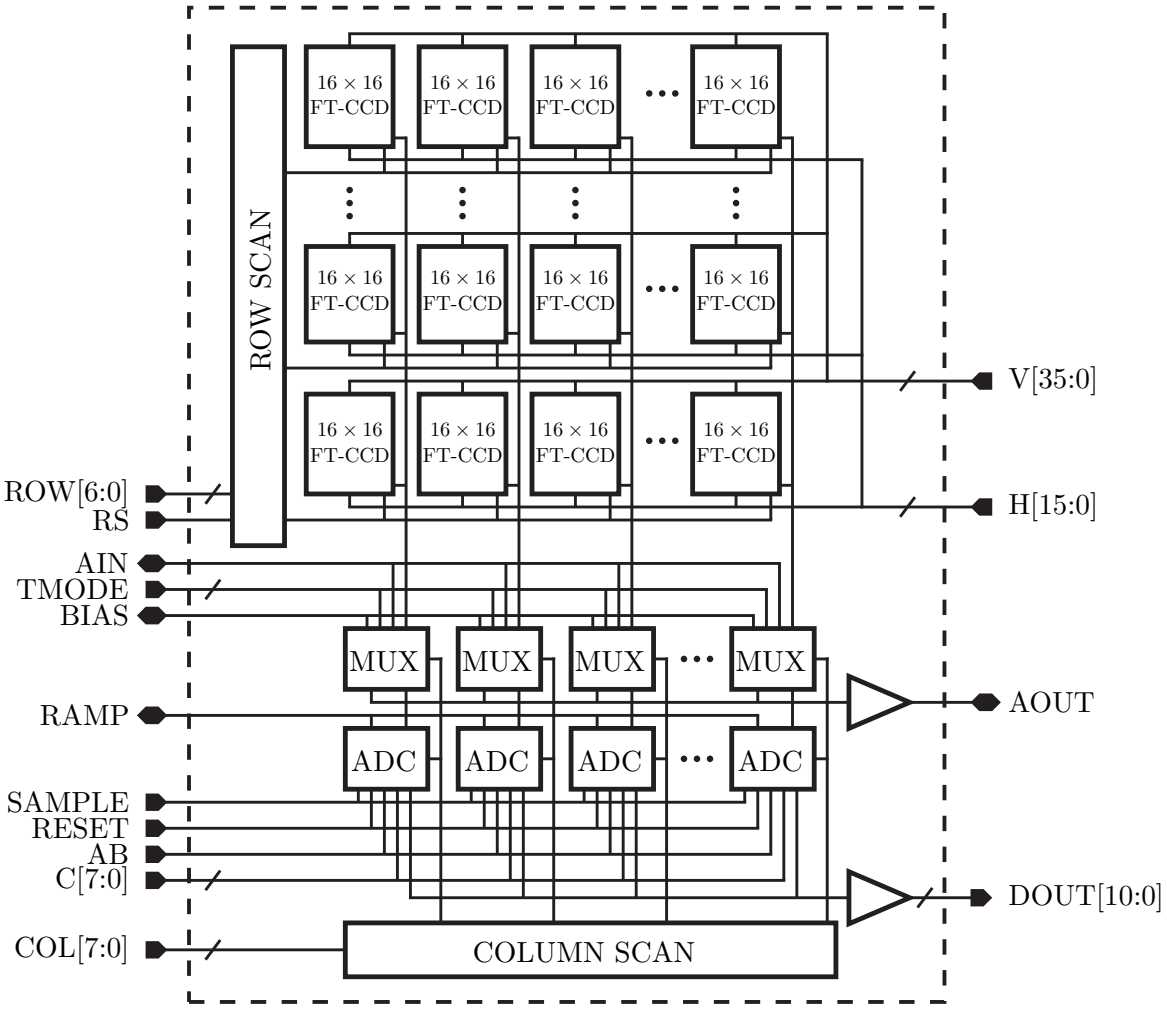



Figure 5.1: Block diagram of multi-aperture image sensor chip.

and to use a channel stop implant instead of a shallow trench. The inputs to the channels at the top of the array are connected to V0 through an n-well implant, which also allows the p-type channel stops to remain isolated from the channels while connecting to ground. Two sides of the H-CCD connect to VP, which is used for fill-and-spill operation, reset of FD node, or as source follower drain supply. The image is captured in interlaced fields. Charge is collected under every other electrode, which allows large potential barriers between them leading to high well capacity. An STI region is used to create isolation between arrays and to serve as the area for contacts to the non-silicided electrodes. The gap width between the polysilicon electrodes is 180nm, which is small enough to implement a single electrode CCD at CMOS compatible voltages.

The layout of the FT-CCD from Diffusion up to Metal 1 is shown in Figure 5.3. A mask is used to block silicide on the polysilicon in order to improve the transmissivity. Metal 1 is used to both cover the frame buffer region and to provide the signal routing to the H-CCD electrodes. The electrodes in the vertical CCD (V-CCD) are constructed from long continuous lines of polysilicon. Since the active area has no silicide, the sheet resistance of the polysilicon is very high. In order to achieve fast transfer times, the time constants for driving the long lines of polysilicon should be minimized. All of the V-CCD electrodes are connected globally, so it is possible to route them either horizontally or vertically. In order to reduce the time constants required to drive the polysilicon lines, vertical connections are made to the polysilicon between subarrays. However, since the space is very tight between neighboring subarrays, we can only connect one polysilicon line to one Metal 1 line at each subarray interval. Therefore, we step the connections to the electrodes at each subarray interval such that after 35 steps all electrodes have been connected. The step interval is 14.4 $\mu$m, which leads to a contact spacing on each electrode of $35 \times 14.4\mu\text{m} = 504\mu\text{m}$.

The layout of the FT-CCD readout circuit from Diffusion up to Metal 1 is shown in Figure 5.4. The circuit is situated in between subarrays such that the fill-and-spill input diffusion node VP is shared with the neighboring drain diffusion of the reset device. Since the fill-and-spill operation is applied only during non-readout times, there is no conflict in using VP to supply the current for the source follower during readout.

### 5.1.1 FT-CCD Operation

The FT-CCD performs snap shot imaging using a global electronic shutter. The capture of a frame can occur simultaneously with the read out of a previous frame. To minimize pixel pitch, ripple transfer operation (as opposed to the more common multi-phase transfer) is

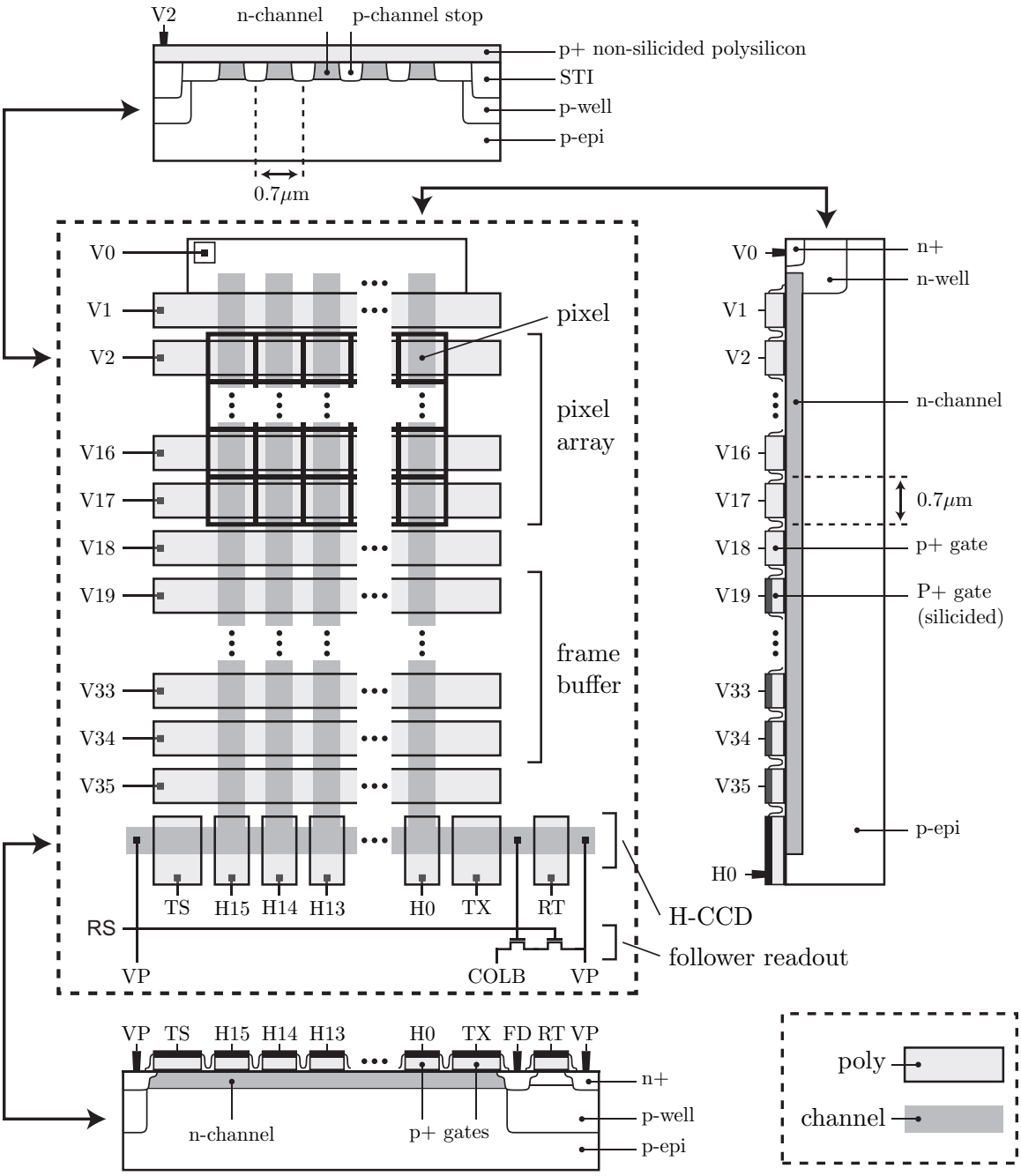



Figure 5.2: FT-CCD schematic and device cross sections.

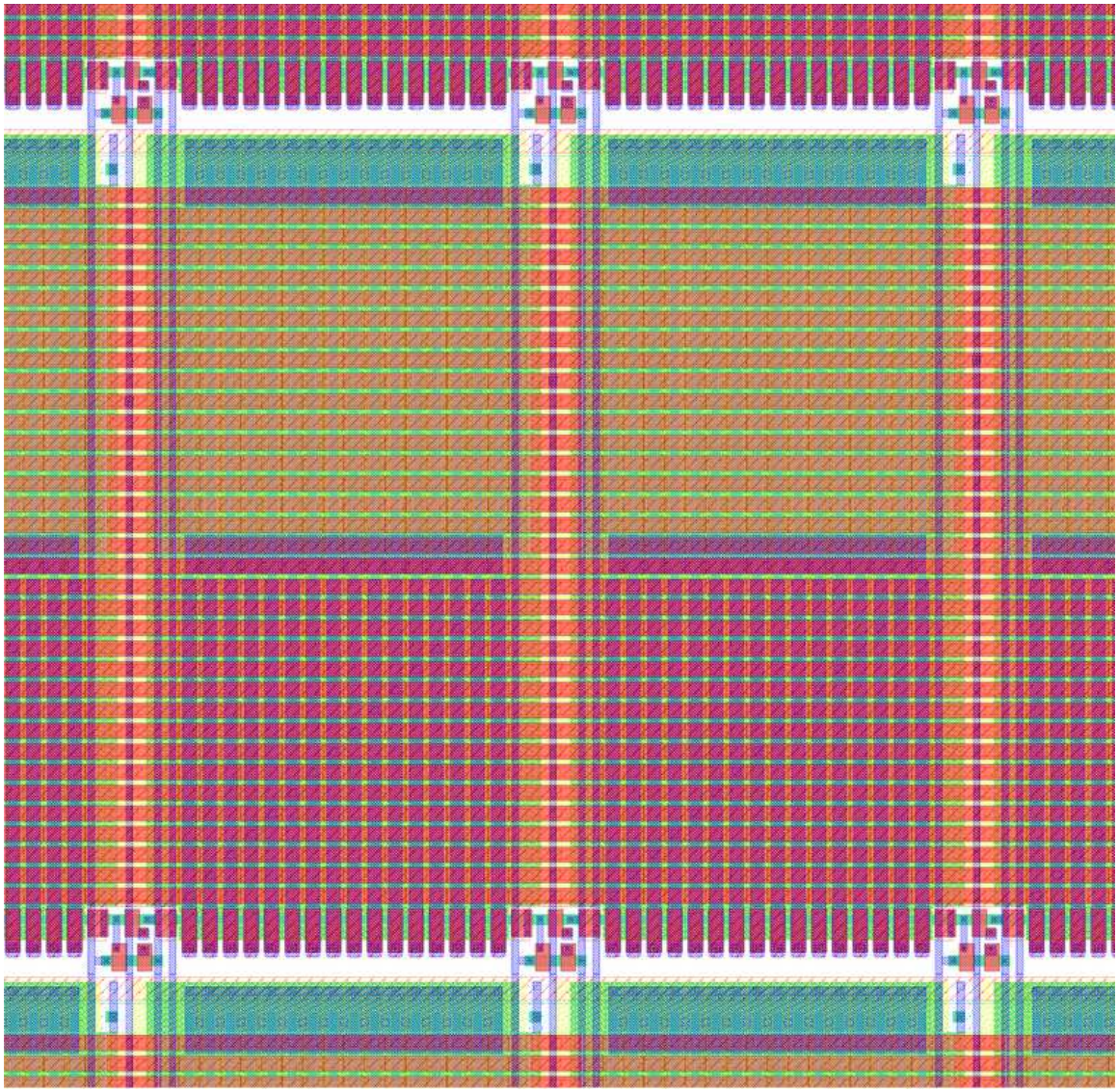


Figure 5.3: CAD layout of two 16×16 FT-CCD subarrays showing diffusion, polysilicon, contacts and Metal 1.

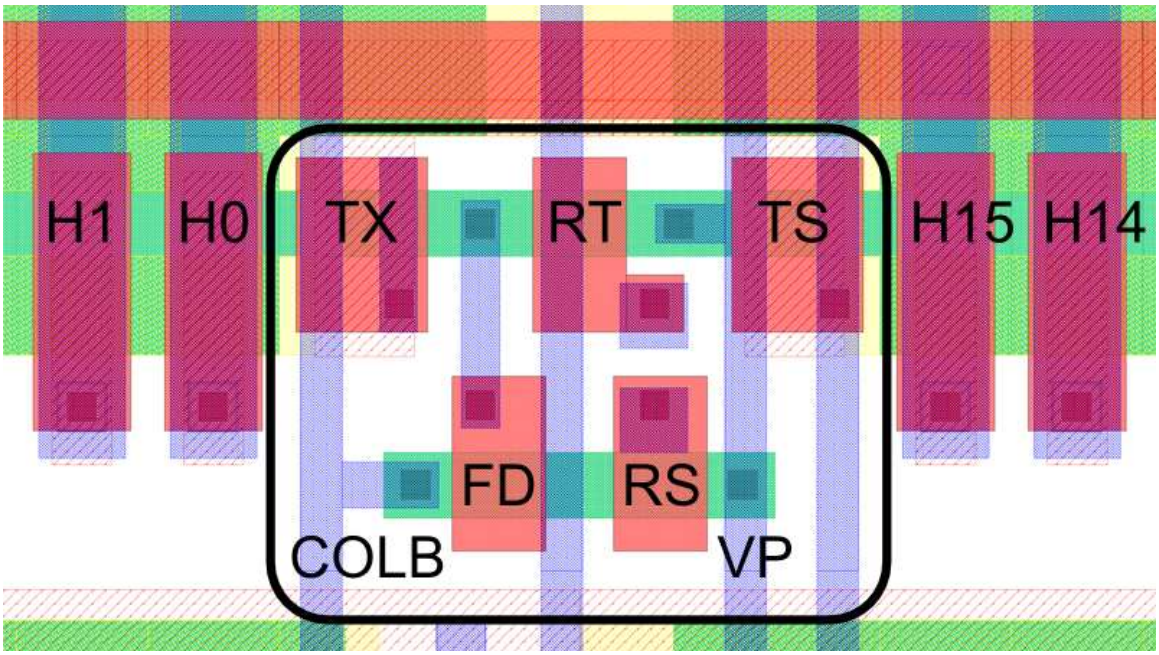



Figure 5.4: CAD layout of CCD readout circuit.

used. An image can be captured by integrating photocharge at each electrode or at every other electrode for interlaced operation. In interlaced operation, the pixel is effectively twice as long as it is wide during each field.

The basic phases of the FT-CCD operation are described with the help of the timing diagram in Figure 5.5. During FLUSH, the CCD pixel arrays are depleted of charge through V[0] by sequencing V[17:1]. During INTEGRATE, the pixel array electrodes are held at an intermediate voltage. At the end of integration, the accumulated charge in the CCD pixel arrays are transferred one row at a time to the frame buffers using ripple charge transfer (TRANSFER). After transferring all of the integrated charge to the frame buffer, FRAME BUFFER READOUT is performed while a new FLUSH cycle is initiated to set the next integration time period or to prevent new charge from flowing into the frame buffer. DIGITAL READOUT at the IO pins is synchronized to FRAME BUFFER READOUT through the ADC interface. A small VBLANK period, where no data is read, is required during the TRANSFER period.

The FRAME BUFFER READOUT sequence consists of a Vertical-to-Horizontal (V-to-H) transfer sequence followed by a Horizontal transfer to the floating diffusion node as shown in Figure 5.6. The goal of the V-to-H transfer is to move just one charge packet at

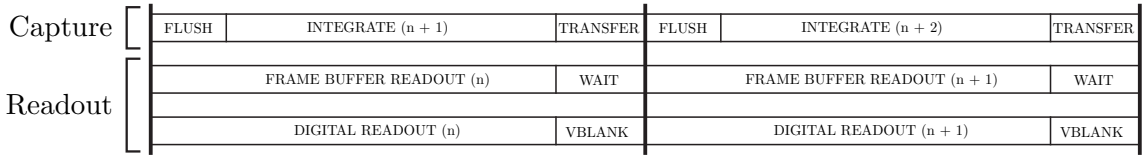


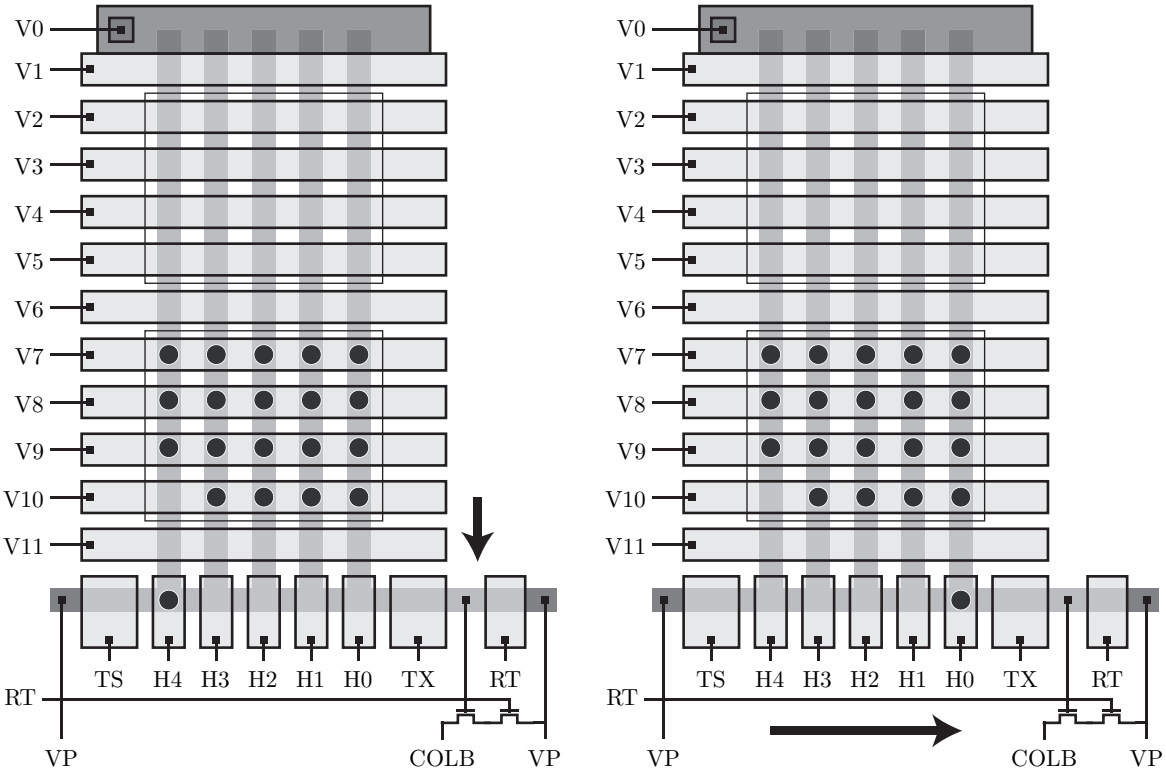

Figure 5.5: Timing diagram for operation of the FT-CCD showing simultaneous capture and readout sequences.

a time into the H-CCD. When the charge packet is shifted along the H-CCD, it is essential to preserve the charge in the other columns. In [49], we described a scheme where the even columns are first transferred to the H-CCD, the H-CCD is drained, and then the odd columns are transferred into the H-CCD. This method has the disadvantage of keeping the charge packets held in the H-CCD for a longer period of time. We find that the defect density in the H-CCD is larger than that of the V-CCD. For this reason, we see higher dark current correlated by columns when this even/odd pixel readout approach is used.

### 5.1.2 FT-CCD Simulation

The main reason we are able to achieve very small pixel size is that we use ripple charge transfer instead of the more conventional multi-phase transfer used in larger pixel count CCDs. A ripple charge transfer requires independent access to each of the electrodes to allow for charge confinement at each electrode. As such, each of the 35 vertical and 16 horizontal electrodes used in our FT-CCD are driven separately.

Process and device simulations were performed with Sentaurus TCAD. A full simulation of the single electrode charge confinement and ripple charge transfer is shown in Figure 5.7. Initially, all electrodes are held at -0.5V, which creates potential wells between every electrode as shown in Figure 5.7a. In this example, we show how charge packets can be placed at V1, V2, V3, and V4. To the best of our knowledge, this is the first CCD to use charge confinement between electrodes, as opposed to under the electrodes. To move one charge packet forward, 3.0V is applied to V4 while 2.0V is applied to the rest of the electrodes in the direction of the charge packet's movement as shown in Figure 5.7b. In Figure 5.7c, the process continues with a ripple transfer, where V5 is brought to 3.0V while V4 is driven back to -0.5V. The charge packet reaches the end of the CCD line as shown in Figure 5.7d. Next, another charge packet is brought forward as shown in Figure 5.7e. Note that it is essential to have an empty well between the packet already transferred and the new one. In a typical CCD this is achieved by having additional electrodes in the pixel.




(a) V-to-H transfer (b) Horizontal transfer

Figure 5.6: FT-CCD readout sequence from the frame buffer.

Here the empty wells are created at the end of each charge transfer. After reaching the state shown in Figure 5.7f, we have to be careful that charge packets do not mix on the next transfer. With the right choice of voltages, the scenario in Figure 5.7g is created, where the left and right sides of electrode V6 are at different potentials, but with a large enough barrier under V6 to prevent charge mixing. Finally the state shown in Figure 5.7h is reached, where two charge packets have moved from one side of the CCD to the other. This is the same way charge is moved during Frame Transfer and along the H-CCD. To increase well capacity, only the even electrodes are set to store charge while the odd electrodes are set to form barriers. A simulation of this setup is shown in Figure 5.8(a). This *interlaced* imaging mode of operation results in a significantly a larger well capacity. The potential profile for the odd field is shown in Figure 5.8(b).

A simulation of the V-to-H-CCD transfer sequence is shown in Figure 5.9. The charge packets in the columns under V34 are shifted to V35 as shown in Figure 5.9b. The horizontal electrodes are initially held at -0.5V to keep all charge under V35. The targeted horizontal electrode is then brought to 3.0V, which forces the targeted column charge to drain into the H-CCD as shown in Figure 5.9d. Potential barriers around the target horizontal electrode are enforced by holding the other electrodes at -0.5V, while the p+ region under the horizontal electrode (shown in the figure) provides isolation along the horizontal axis. Next, V34 is brought to 3.0V and a partial transfer occurs at the other columns from V35 to V34 as shown in Figure 5.9e. A full charge transfer is now achieved in both directions by slowly dropping V35 to -0.5V as shown in Figure 5.9f. This transfer mechanism relies on the condition that the fringing field from horizontal electrodes remain larger than that from V10, and that V11 provides a sufficient barrier. Once charge is completely transferred to the horizontal CCD, V34 is set close to 2.0V to ensure that all the non-targeted column charge is efficiently passed backwards. Next, the charge in the H-CCD is ripple shifted to the floating diffusion node where it is buffered and double sampled using a source follower circuit. This procedure is repeated for the remaining charge in the other columns confined by the V34 electrode.

### 5.1.3 FT-CCD Readout Circuit

The local readout circuit for the FT-CCD is shown in Figure 5.10a. A long TX electrode provides sufficient isolation from the H-CCD to allow the FD node to be driven into reset while the other charge packets remain undisturbed in the H-CCD register. Note that the TX electrode can be eliminated from the design if we choose to move just one packet at a time

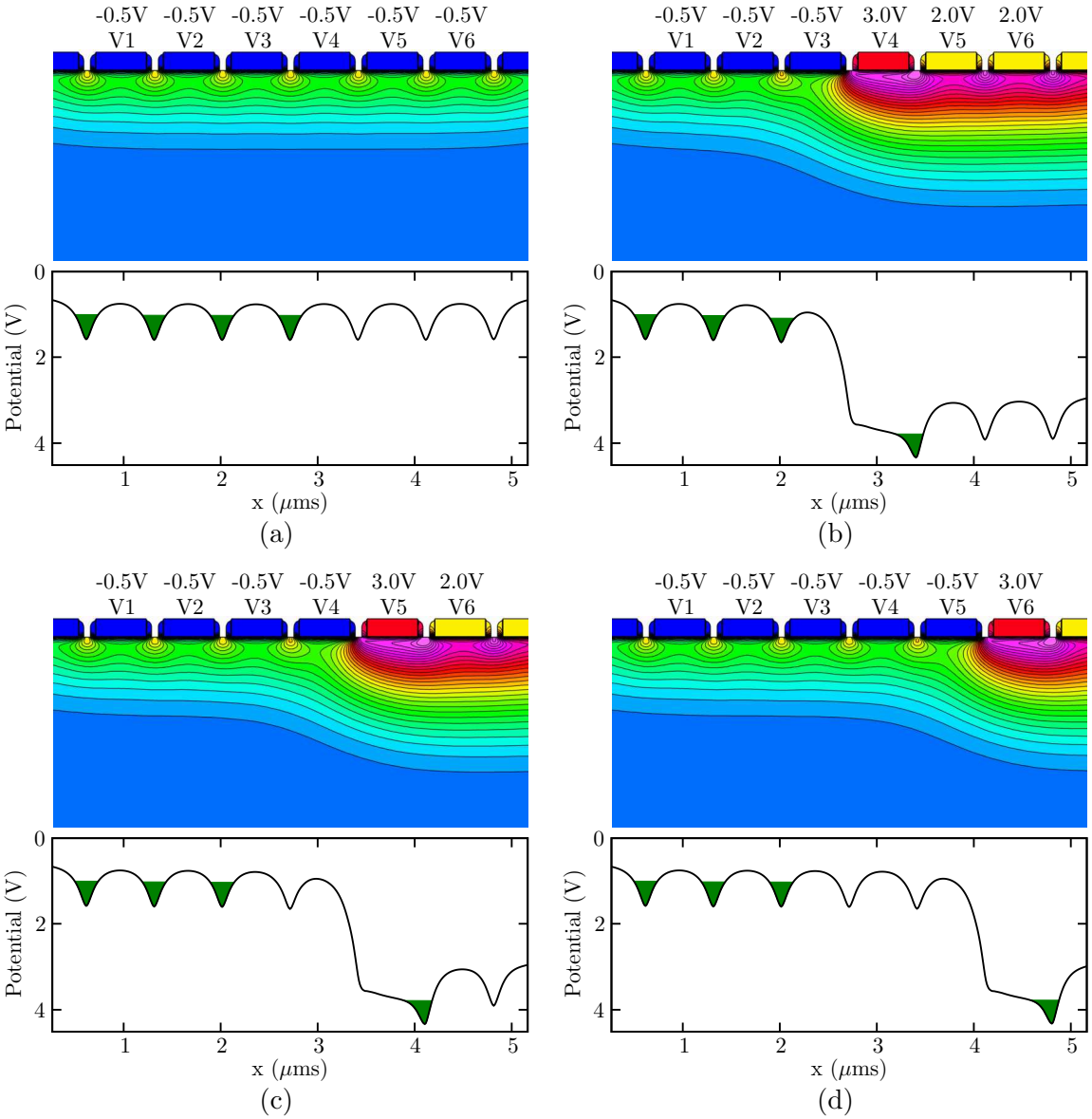



Figure 5.7: Electrostatic potential diagrams from the simulation of CCDs with single electrode charge confinement and ripple charge transfer.

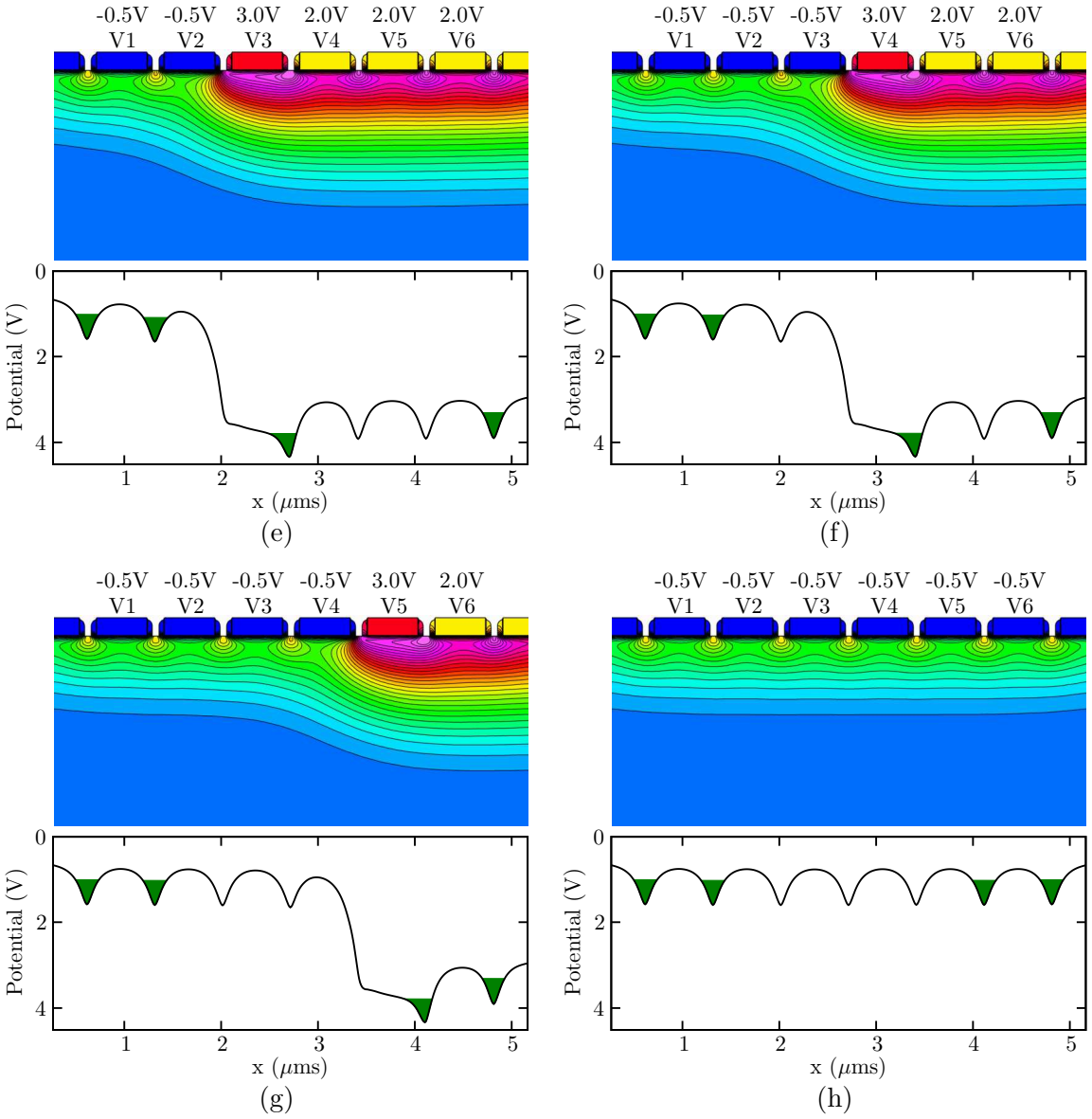
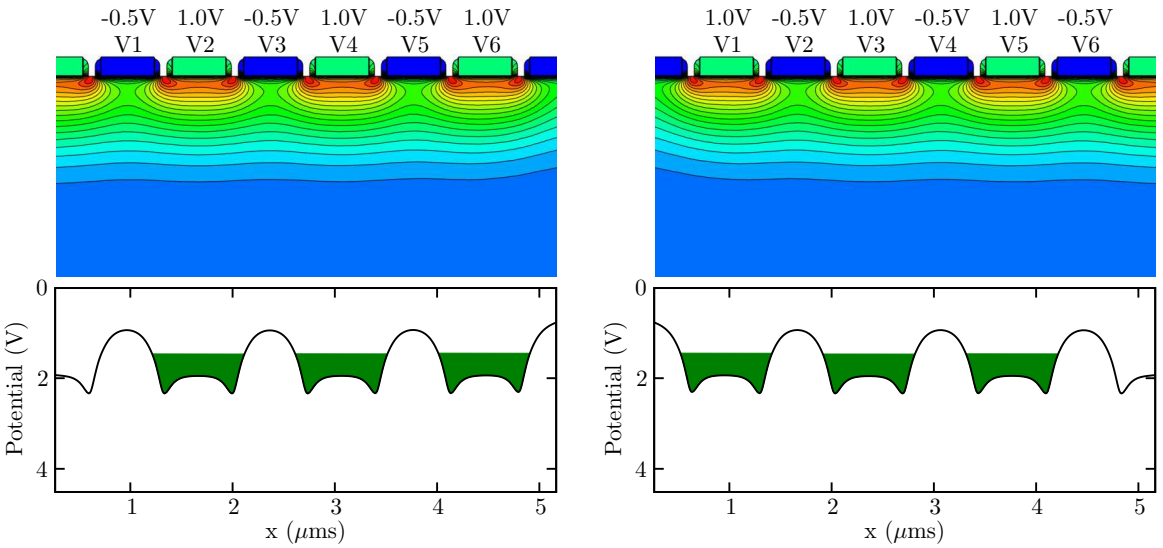



Figure 5.7: Electrostatic potential diagrams from the simulation of CCDs with single electrode charge confinement and ripple charge transfer (cont.).




(a) Even field (b) Odd field

Figure 5.8: Electrostatic potential diagrams under interlaced frame operation showing increased size of the potential wells.

during the V-to-H transfer. The readout circuit consists of a floating diffusion with reset and row select. To achieve low noise, we perform Correlated Double Sampling (CDS [50]) on the floating diffusion node. First, FD is reset by bringing RT high, which turns on the transistor $M_1$. The voltage on VP is sufficiently low with respect to RT such that FD receives a hard reset. RT is then driven low followed by RS driven high. Drain current is now provided through $M_2$ for the source follower $M_3$, which is biased at COLB by $I_0$. The value that settles on COLB represents the reset level in addition to the sampled noise. Next, TX is brought high and the charge packet at H0 is shifted to the FD node, which decreases its voltage. The value that settles on COLB is the signal subtracted from the reset and sampled noise. Taking the difference between the 2 samples leaves just the signal without the reset and noise.

To deplete the CCD channel and achieve a full charge transfer, we must keep the potential at FD higher than the channel potential under the depletion condition. For this reason, we boost the FD node potential during the transfer operation. To describe the readout operation in detail, consider the circuit with the most significant parasitic capacitors shown in Figure 5.10a and the corresponding timing diagram in Figure 5.10b. The operation shows the 4 main phases previously described but now with more detail including the effect of the parasitic capacitors. For example, the VP line is switched between a high potential

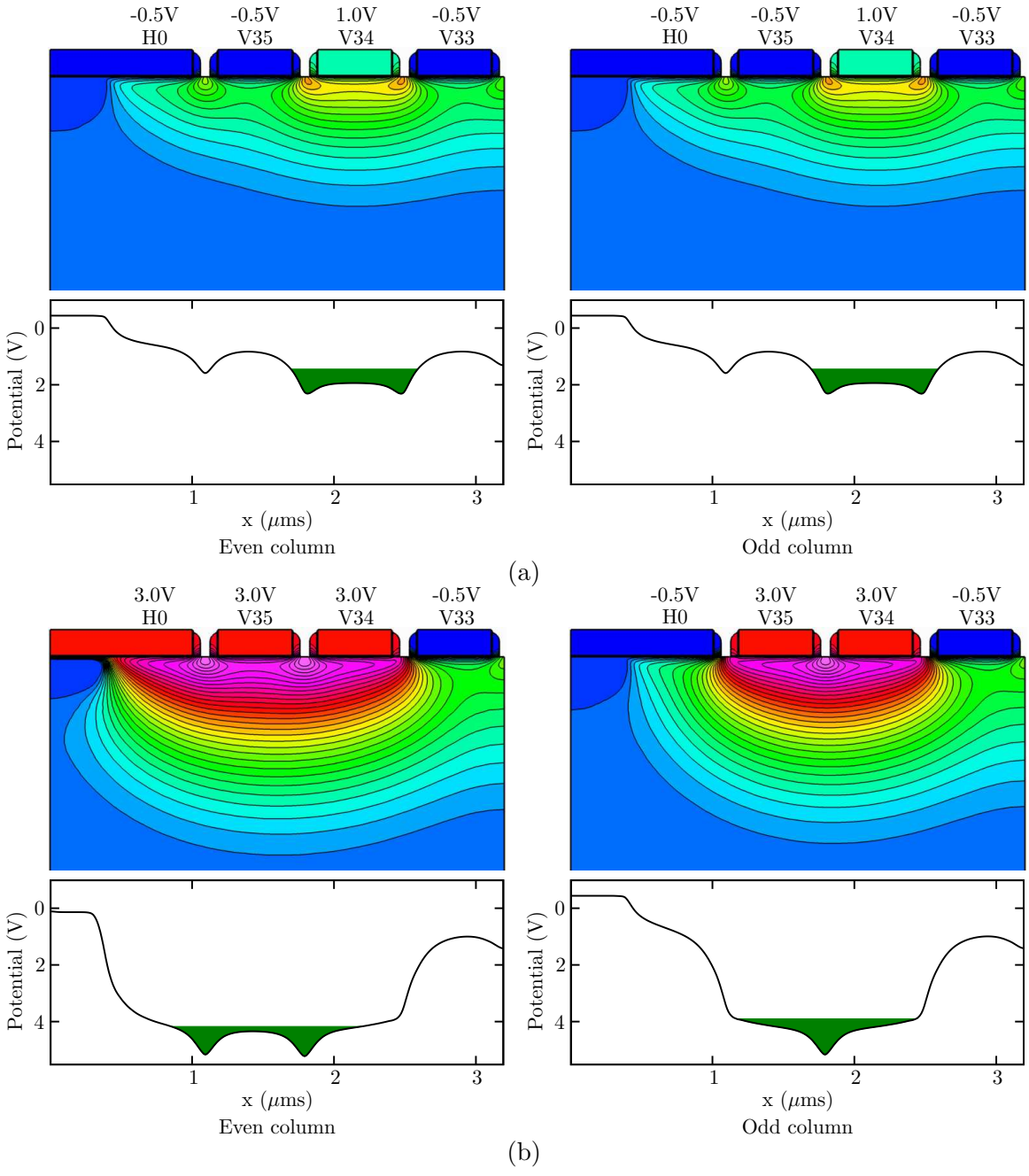
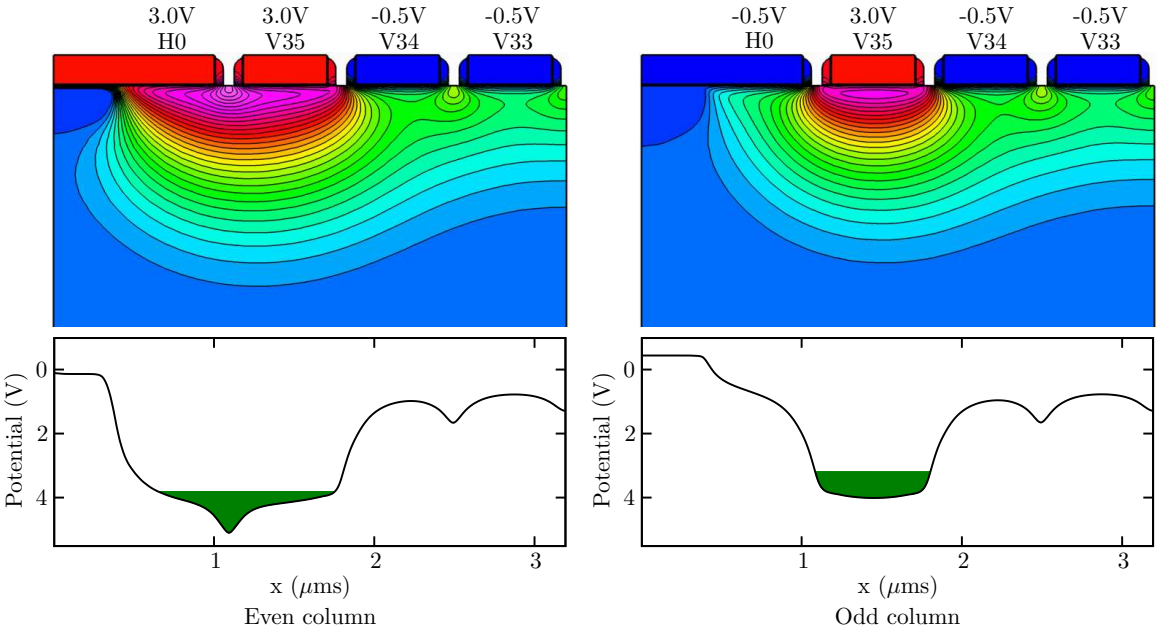
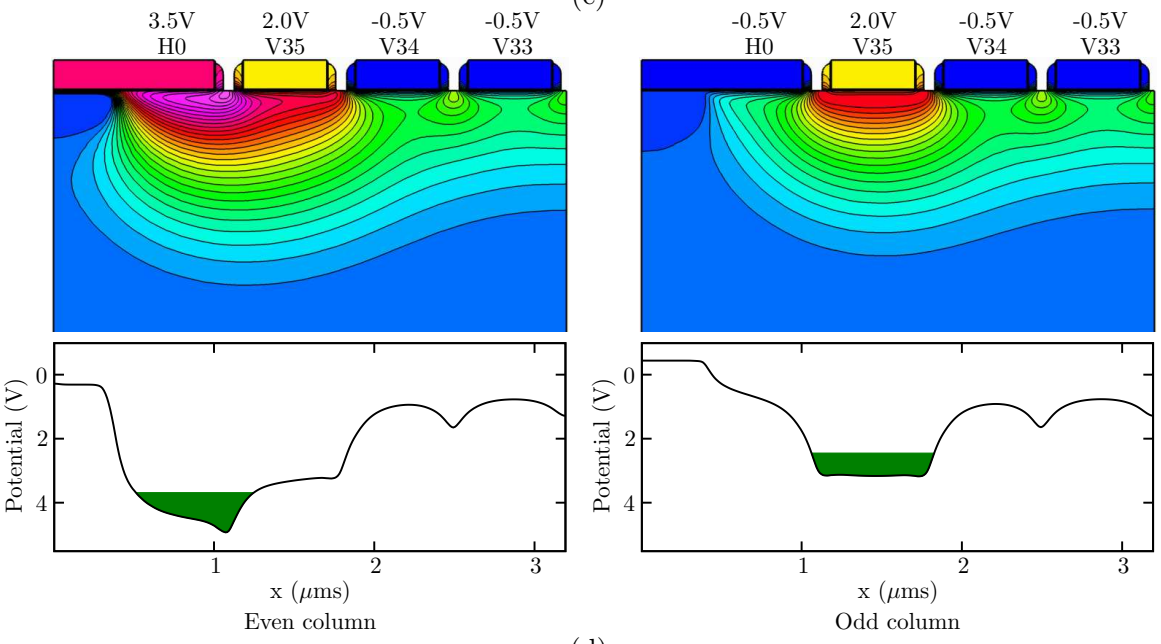



Figure 5.9: Details of the vertical to horizontal transfer. (a) Charge is initially confined under V34 for both columns. (b) Charge packets under V34 are shifted to V35. The even horizontal electrode is brought high, which forces the even column charge to drain into the H-CCD.




(c)




(d)

Figure 5.9: Details of the vertical to horizontal transfer (cont). (c) V34 is brought low to complete the transfer to V35. (d) V35 is brought to a low enough potential that charge is shifted under H0.

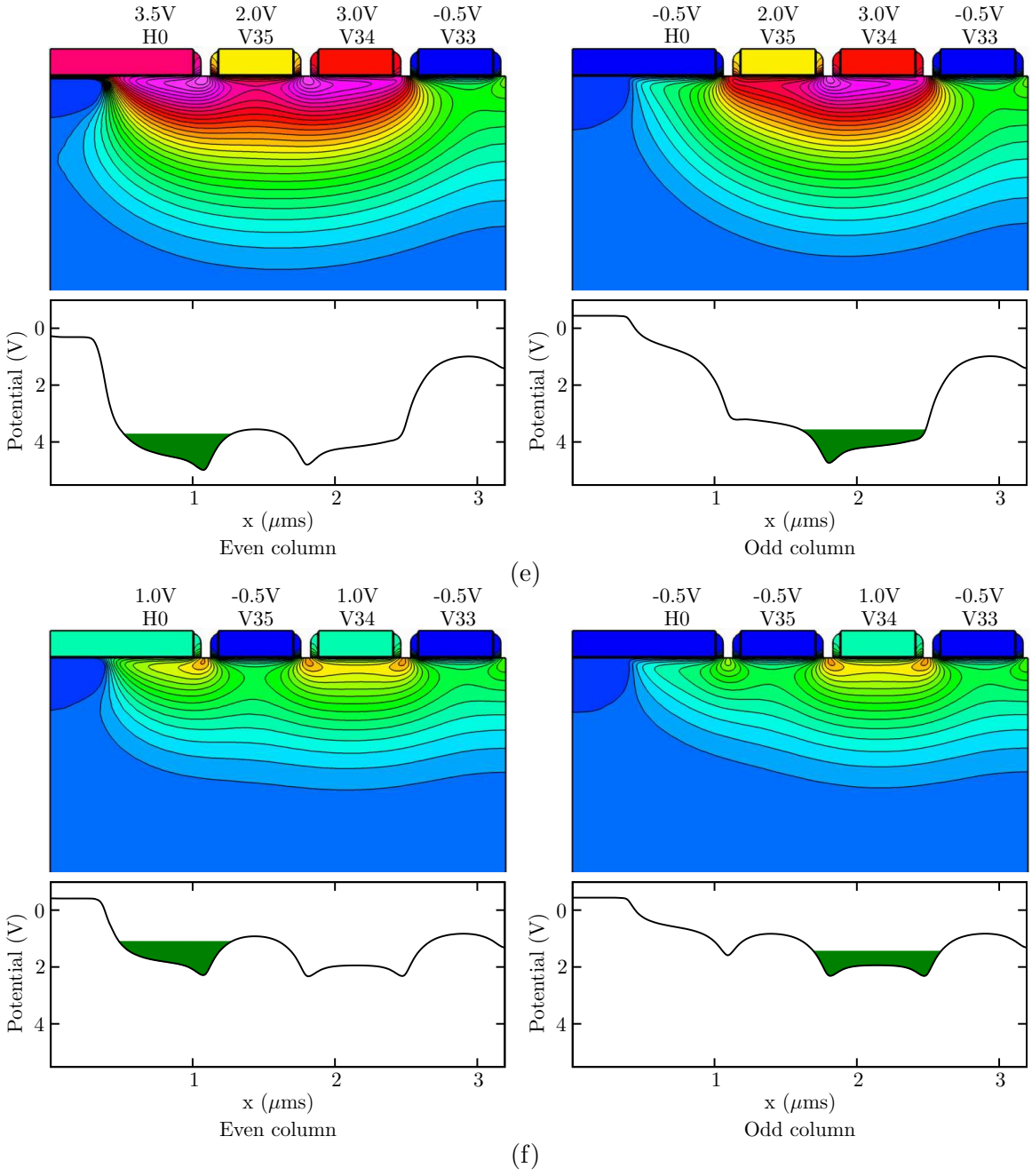



Figure 5.9: Details of the vertical to horizontal transfer (cont). (e) V34 is brought high and a partial transfer occurs at the odd columns from V35 to V34. (f) Final potential profile for complete charge transfer.

and a lower potential before the RT gate is driven high. Coupling through $C_3$ causes the FD node to dip down slightly as shown. Driving the RT gate high resets FD to VP but the final reset value becomes lower after the RT gate is released low due to the feedthrough from the gate-source overlap capacitance $C_1$. When FD is sampled during the Sample R phase, driving RS high causes the FD node to boost higher due to coupling from $C_4$, which sees a nearly full range terminal switching. In addition, the gate-source overlap capacitor represented by $C_5$ discharges into FD causing it to rise even higher. We can take advantage of this effect by driving COLB high during the transfer phase. We create significant FD boosting during the Transfer operation as shown. The final boosted value is controlled by the selection of the switched VP voltage during Reset so it can be finely tuned. Under the case of significant boosting, it is better to drive TX high before boosting in order to incur less stress on the gate oxide (The timing diagram shows boosting occurring first so that the charge transfer is easier to see). Finally, the TX gate is released and RS is driven high for the second sample. The diagram shows the signal level as it appears at FD and COLB.

## 5.2 Chip Operation

The basic timing diagram for 2 rows of the chip is shown in Figure 5.11. The diagram is divided into 6 macro operations for simplicity. A representation of the chip with a 2×2 array of apertures, each with a 2×2 pixel FT-CCD is shown in Figure 5.12. The readout sequence begins after the integration period shown in Figure 5.12(a). The active CCD area contains 4 charge packets, while the frame buffer (shown in gray) has been fully depleted of charge. The FT-CCDs have their own local readout circuits, which are configured by columns each connected to an ADC. All of the CCD operations are global. Therefore, the start and end of the integration period is the same for all subarrays. At the end of the integration period, the charge is transferred to the frame buffer (see Figure 5.12(b)) as described in Section 5.1.1. To begin readout, the reset value for the first pixel of every subarray is first acquired by bringing RT high for every row (see Figure 5.12(c)). Next, the row select signal is applied to RS[0] via the row decoder and the value at each column ADC is sampled (see Figure 5.12(d)). Next, row select is applied to RS[1] and the value at each column ADC is sampled so that reset values from all subarrays are read out (see Figure 5.12(e)). After a charge packet is transferred from the V-CCD to the H-CCD, TX is applied globally to all rows and the charge is transferred to the floating diffusion (see Figure 5.12(f)). Next, row select is applied to RS[0] and the values are sampled at the ADCs (see Figure 5.12(g)). Next, row select

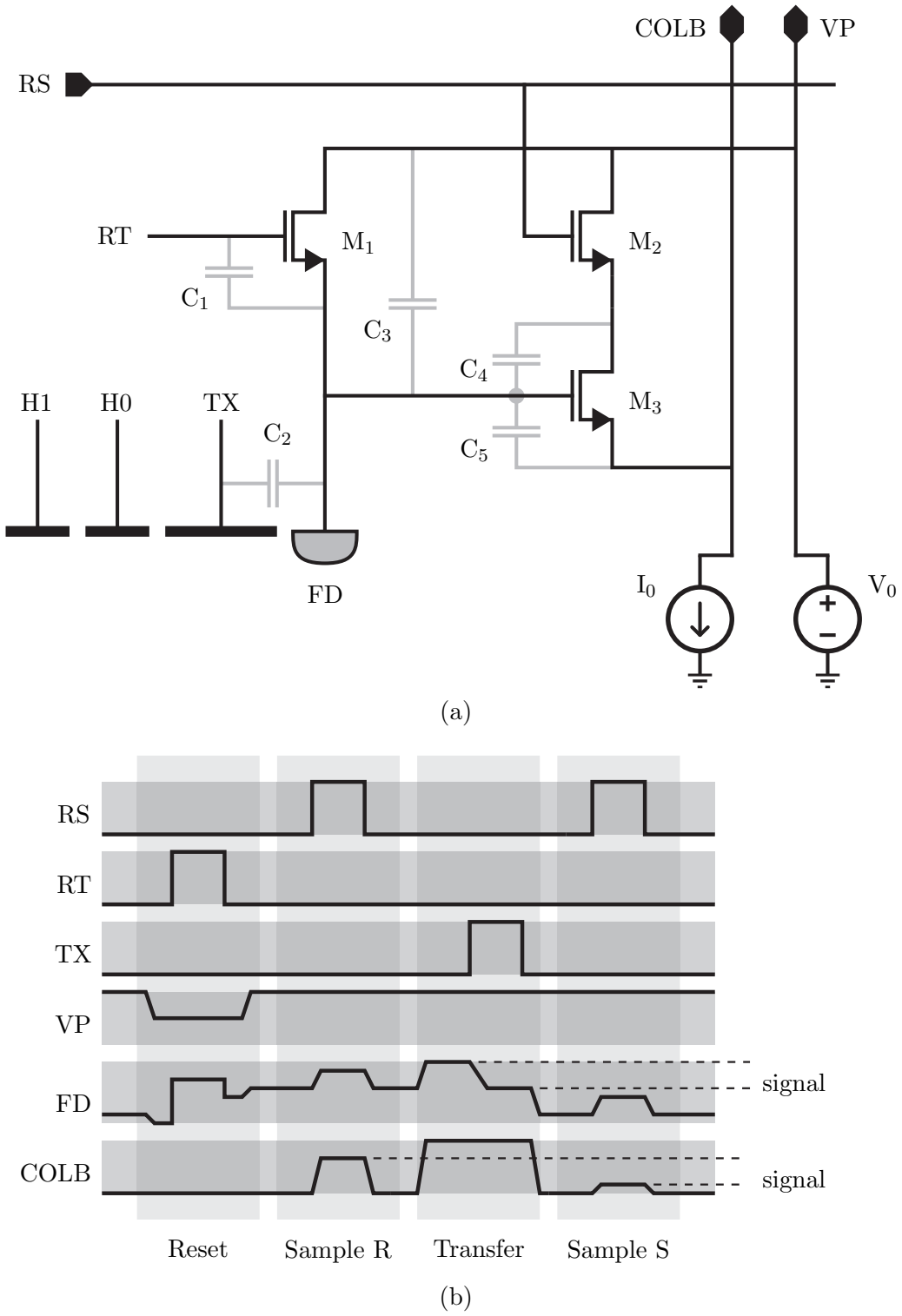



Figure 5.10: (a) Local CCD readout circuit including the most significant parasitic capacitors, and (b) the corresponding timing diagram.

is applied to RS[1] and the pixel values are sampled at the ADCs (see Figure 5.12(h)). After all reset values and signal values are read out, CDS is performed off-chip in software by subtracting each reset value from its corresponding pixel value. The ADCs are double buffered so that the pixels in a row of subarrays can be read out while the previously converted pixel data is read out from the ADCs. After the last pixel is read out, the floating diffusion is reset by bringing RT high for all rows (see Figure 5.12(i)) and the readout sequence is repeated until all pixels are converted. This operation allows us to keep all signals in the image sensor global except for RS. This greatly simplifies the row decoder operation and allows for global shutter operation, which is preferred to the more common rolling shutter. The cost of implementing the image sensor in this way is that row buffers are required to realign the pixel data in order to perform digital CDS. The number of row buffers required is equal to the number of aperture rows in the sensor. Furthermore, one pixel from each aperture is read out before the next pixel is available. A frame buffer is required in order to obtain all the pixels from one aperture. There are a number of other approaches to the global readout architecture. For example, adding the TX signal to the row decoder would allow sequential sampling of the reset level and the signal level such that the digital CDS can be performed immediately at each row, thereby eliminating the need for row buffers. This approach requires storing charge in the H-CCD at electrode H0 for a period of time that depends on its row position in the pixel array. This requires the dark current generated in the H-CCD to be kept low. Another option is to include all of the H-CCD signals and half of the V-CCD signals into the row decoder. This way all pixels for all apertures in a row can be readout at once, which may be preferable. We chose not to implement this readout architecture because we anticipated the need for using complex waveforms, including negative voltages, to experiment with the CCD operation, and there was no advantage to dealing with raster scan data in our system.

## 5.3 ADC Design and Operation

The column ADC circuit and timing diagram are shown in Figure 5.13. A 10-bit single-slope architecture with off-chip ramp (via 14-bit DAC) is used for flexible operation. Conversion begins by resetting the keeper cell. As the COLUMN value settles, SAMPLE is strobed while the bus signal C cycles through the gray code values, which are non-uniformly spaced to account for shot noise level. Once RAMP exceeds COLUMN, the code on C is latched into the A buffer. The keeper stays latched until the beginning of the next conversion

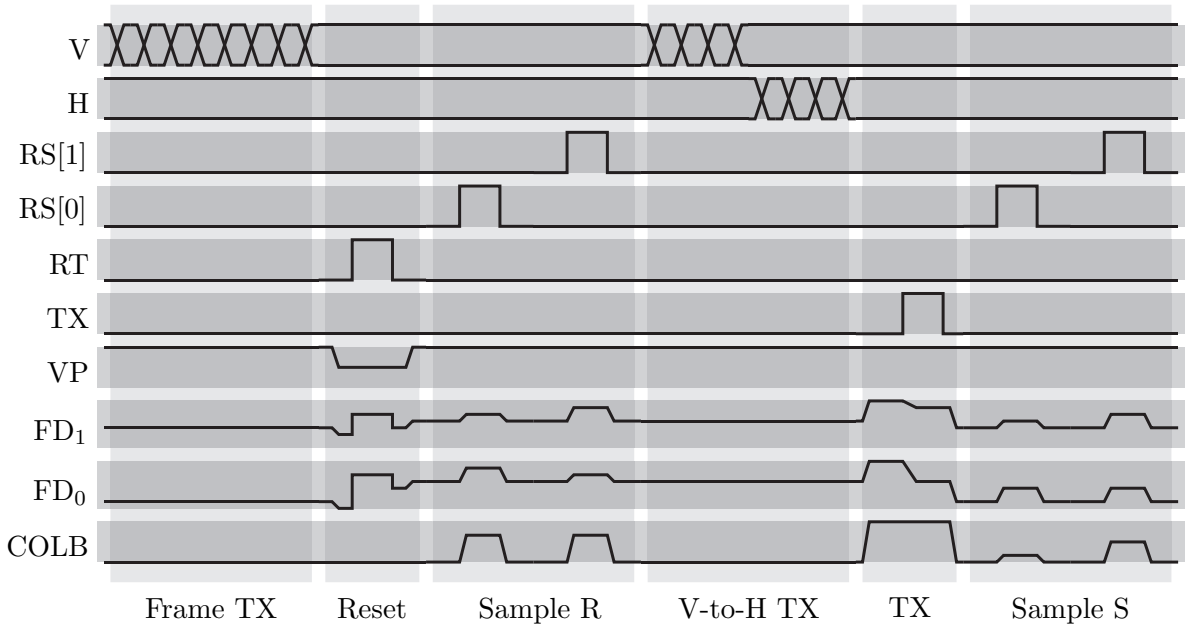
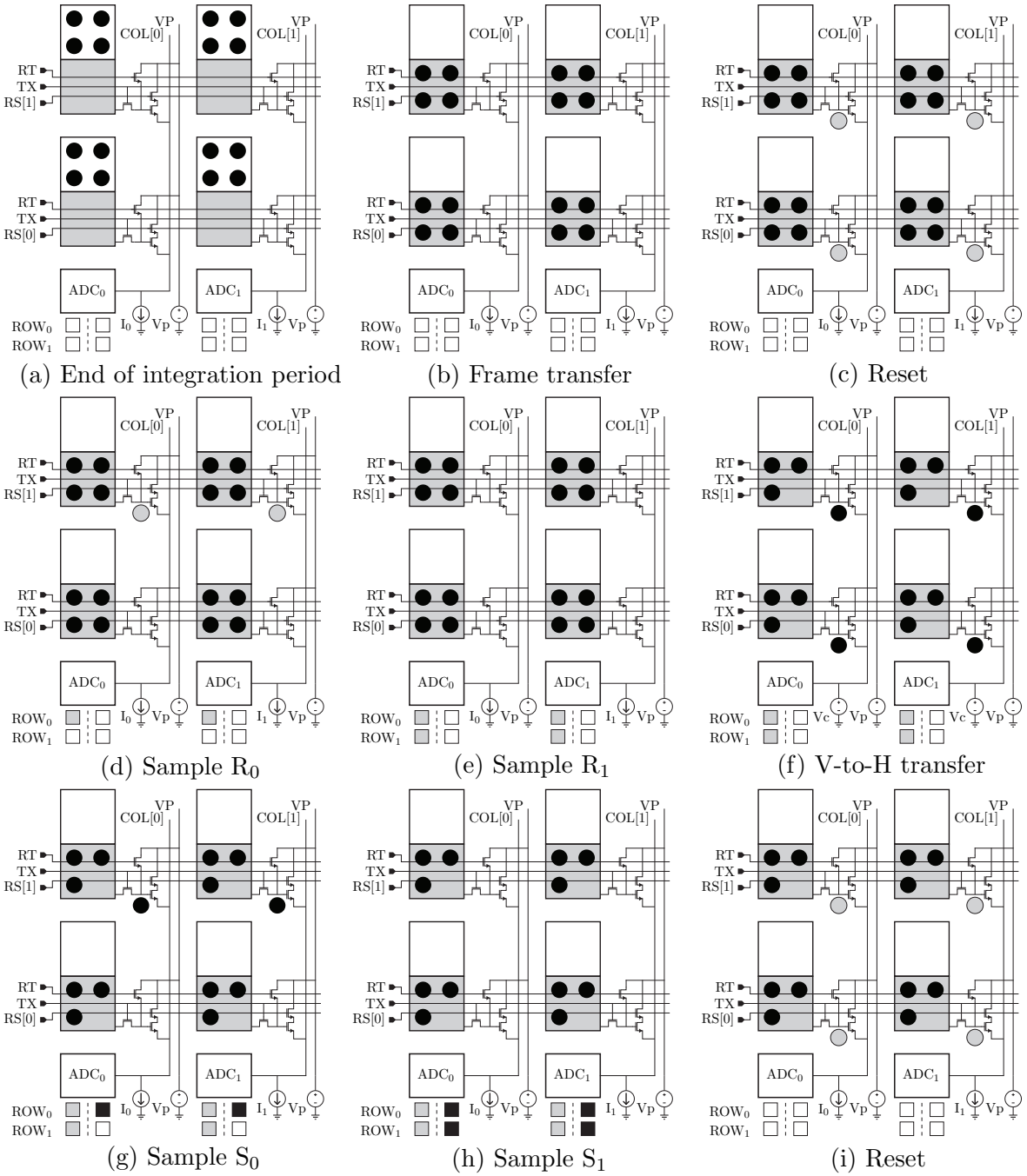



Figure 5.11: Basic chip timing diagram for 2 rows.

where buffer B is used to store C while buffer A is read out. The comparator is capable of operating at 200MSPS, which is required for 15fps operation. It consists of a diff-pair followed by a regenerative latch. Column gain is achieved by reducing the voltage range of the ramp. A resolution of 10 bits at 1V (double sampled pixel) is achieved with 1$\mu$A bias current. The diff-pair transistors, with a W/L ratio of 6, are kept in weak inversion. The comparator bias current can be increased for higher bandwidth. The regenerative portion of the comparator and the memory buffers are implemented with 1V transistors, while the diff-pair is implemented with 3V transistors, allowing for simple translation between power domains.

## 5.4 Characterization Results

The sensor was fabricated in a 0.11$\mu$m CMOS process with modifications to implement the buried channel CCD. A photomicrograph of the fabricated 3.0×2.9 mm$^2$ chip is shown in Figure 5.15. Local optics were not integrated on this chip so all image testing is performed with standard focal plane imaging. The die was packaged in a ceramic 120 pin grid array. A test board, controlled by an FPGA, was built with a DAC for each CCD signal to provide flexibility in the generation of the CCD waveforms. The data is read into a DRAM and




(a) End of integration period (b) Frame transfer (c) Reset

(d) Sample $R_0$ (e) Sample $R_1$ (f) V-to-H transfer

(g) Sample $S_0$ (h) Sample $S_1$ (i) Reset

Figure 5.12: Chip operation showing a complete cycle for reading one pixel from every subarray. The boxes under the ADC represent the digital values for the reset levels (gray) and the signal values (black).

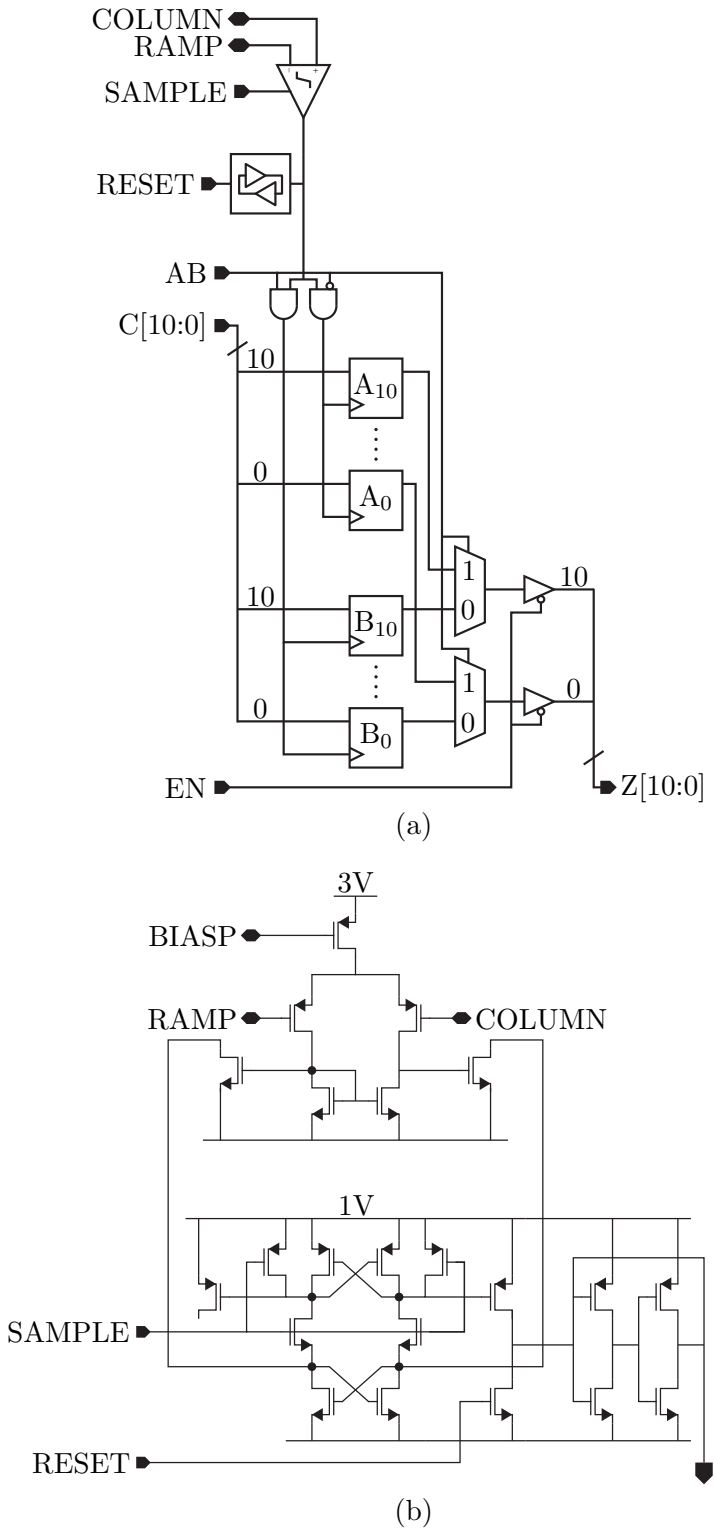



Figure 5.13: (a) Ramp-based column ADC design. (b) Schematic showing transistor level design of comparator and keeper cell.

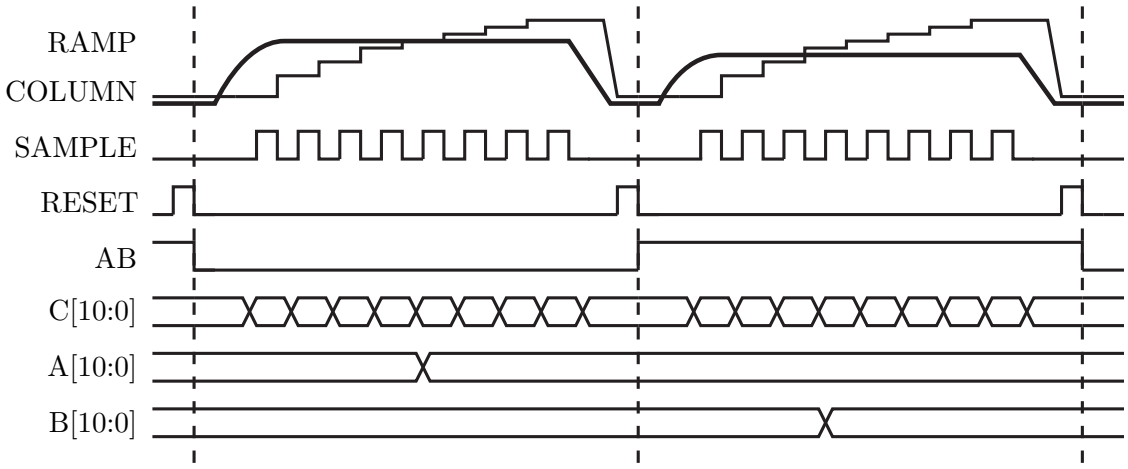


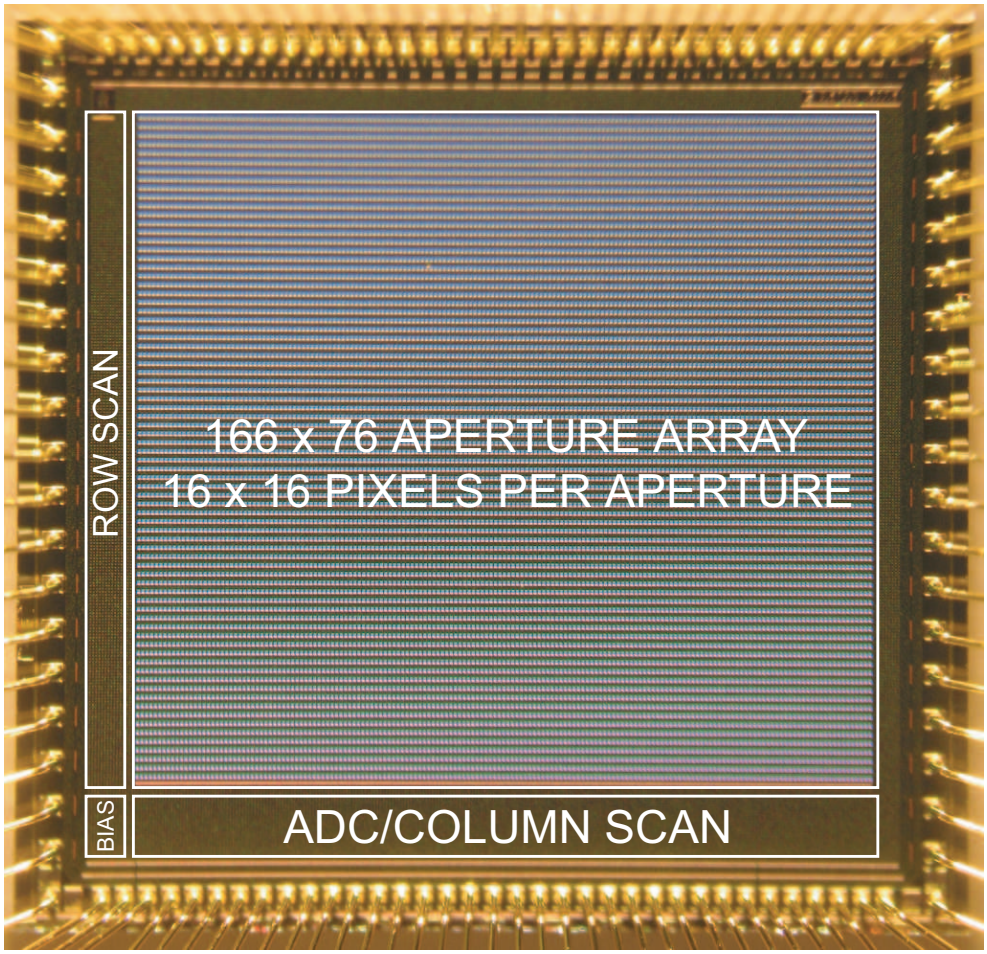

Figure 5.14: ADC timing diagram for 2 conversion cycles.

sent to a PC via a USB interface. The chip is fully functional and has been electrically and optically characterized. All hardware is shown if Figure 5.24.

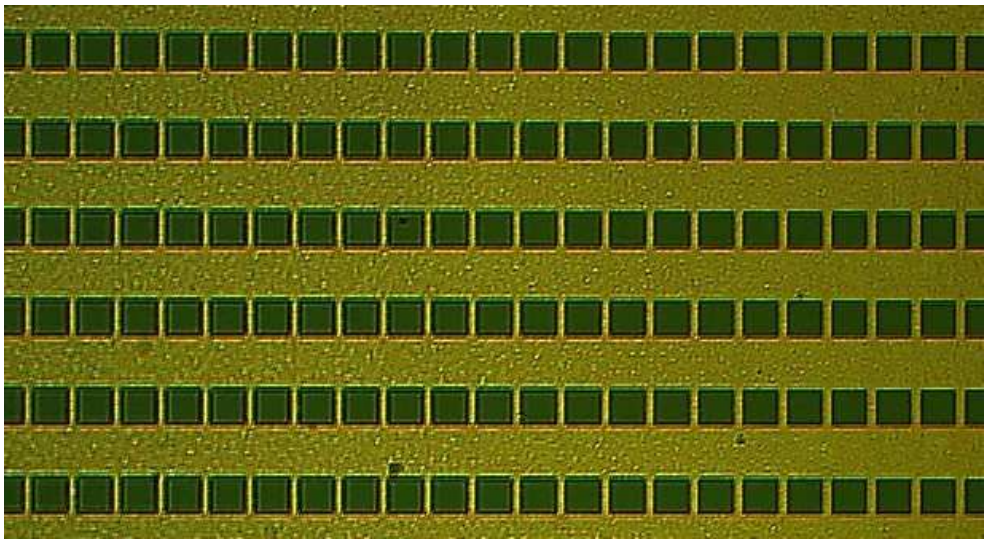
### 5.4.1 Electrical Characterization

The CCDs are electrically tested with the fill-and-spill circuit to verify charge confinement at the 0.7$\mu$m pixel pitch. Charge transfer efficiency (CTE) of 99.9% is measured by moving charge packets through all the columns of the vertical array via the H-CCD. Noise from the fill-and-spill operation is removed by averaging samples and the total transfer times are kept short to minimize corruption from dark current.

The ADC is tested and characterized independent of the pixel array by configuring the MUX for external input via AIN (see Figure 5.1). Two samples are taken for each effective ADC cycle. The first sample is held at 2V, while the second sample is ramped from 1V to 2V to test the ADC linearity and noise over the 1V range. The measured ADC linearity for one typical column is shown in Figure 5.16a. Both the peak-to-peak DNL and INL are kept to about 1/2 LSB on a 10-bit scale. In imaging applications, it is essential to keep column non-uniformity very low. Therefore, the ADCs need to maintain low INL and DNL across all columns in the array. In Figure 5.16b, the peak-to-peak INL and DNL are plotted across all columns showing variation of less than 0.6 LSB among them. Both the fixed pattern noise (FPN) and the temporal noise (TN) of all ADCs in the column array are shown in Figure 5.16c. Due to the digital CDS used in this design, we have an extremely low column FPN of 0.044 LSB RMS and 0.06 LSB max. The temporal noise is also acceptable




(a)



(b)

Figure 5.15: (a) Chip micrograph and (b) magnified view of aperture array.

at 0.26 LSB (254$\mu$V) RMS and 0.6 LSB (586$\mu$V) max, considering the noise floor of the pixel is 825$\mu$V RMS. Table 5.1 provides a summary of the chip-level performance, which is largely determined by the ADC.

### 5.4.2 Optical Characterization

To obtain acceptable results for well capacity, peak SNR, and dynamic range, all measurements are performed using the interlaced mode described in Section 5.1.2. Many key CCD performance parameters can be measured from the photon transfer curve, which is a plot of the measured RMS noise versus the mean of the signal in electrons. The measured photon transfer curve for the FT-CCD subarray is shown in Figure 5.17. From the curve, we find that the noise floor, i.e., the noise at very low signal, is around 5e-. The sensor conversion gain, which is derived from the shot noise limited regime, is 165$\mu$V/e-. The noise near the top of the curve begins to decrease, signifying a full well of 3500e-. The expected decrease in noise is due to charge mixing. The variation in the gain of each pixel is measured as Photo Response Non-Uniformity (PRNU) of 0.02 RMS.

The dark current is measured over several integration times to remove any offsets. The histogram for the dark current over all pixels in image sensor is shown in Figure 5.18a. The mean dark current of 33e-/sec with Dark Signal Non-Uniformity (DSNU) of 0.35 (sigma/mean) is within the expected range for this device. The DSNU is also calculated independently for each 16×16 subarray in the image sensor and shown in Figure 5.18b. Using additional implants to increase the hole concentration near the surface may help reduce the dark current and non-uniformity.

The measured quantum efficiency is shown in Figure 5.19. Despite the use of polysilicon electrodes, the blue response is acceptable. This is due to the thin (130nm) polysilicon layer and the open space in between each electrode. Table 5.2 provides a summary of the measured sensor imaging characteristics.

A sample image acquired in a standard focal plane configuration using a fixed focus F/2.4 lens is shown in Figure 5.20. Since local optics is not implemented on this chip, the images are directly projected to the subarrays. A black grid is inserted into the image data to show what the image looks like at the sensor. The image in the upper right shows the average value at each aperture. The image in the lower right shows a magnified portion of the image to show detail within each subarray.

Sample images from a single subarray are shown in Figure 5.21. The electrical image in Figure 5.21a is generated using the fill-and-spill circuit to input the charge pattern into

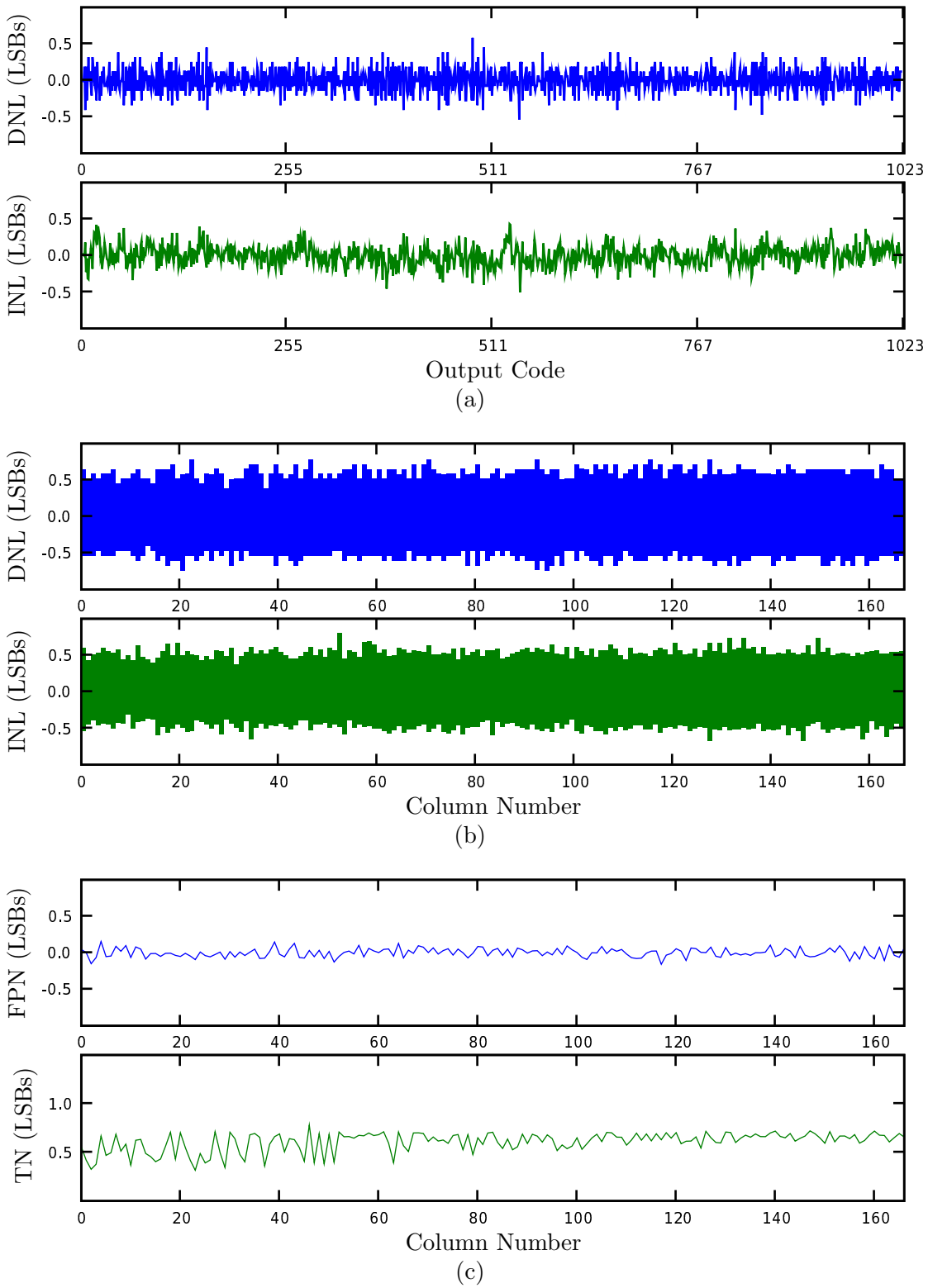



Figure 5.16: Measured ADC performance. (a) Linearity for one column. (b) Minimum and maximum DNL and INL for all 166 columns. (c) Fixed pattern noise (FPN) and temporal noise (TN) for all 166 columns.

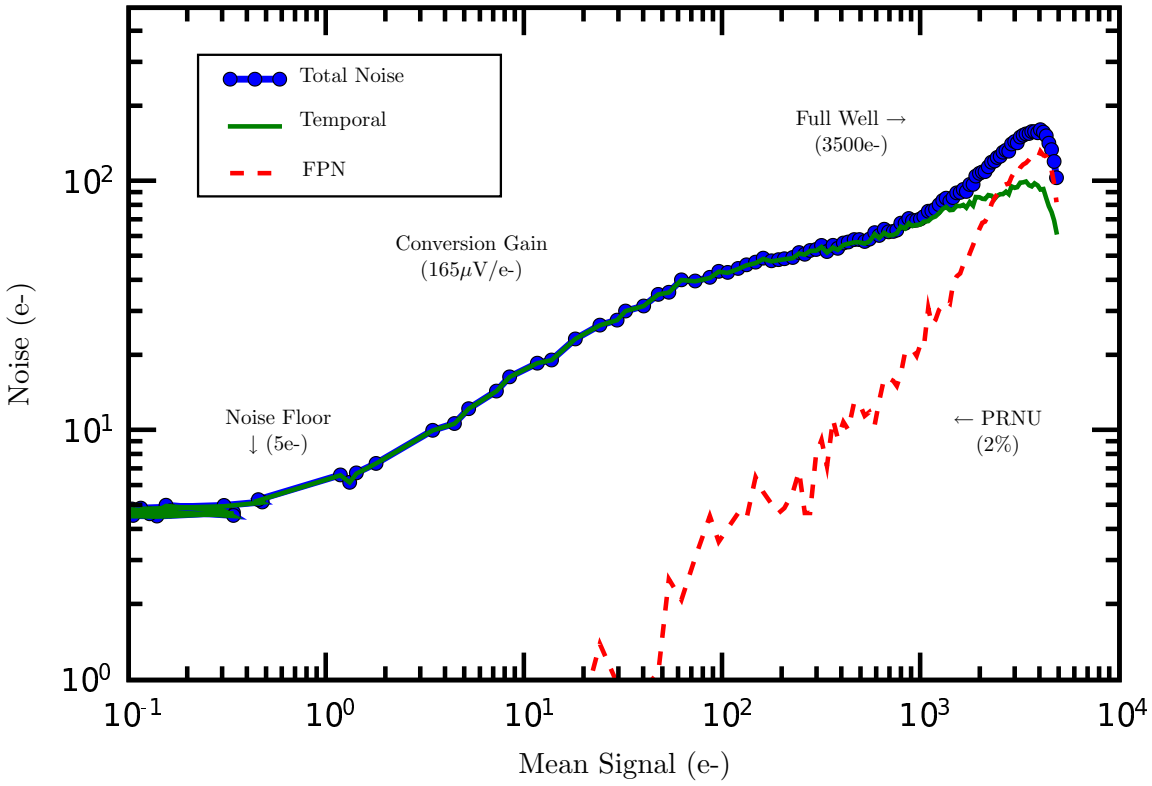
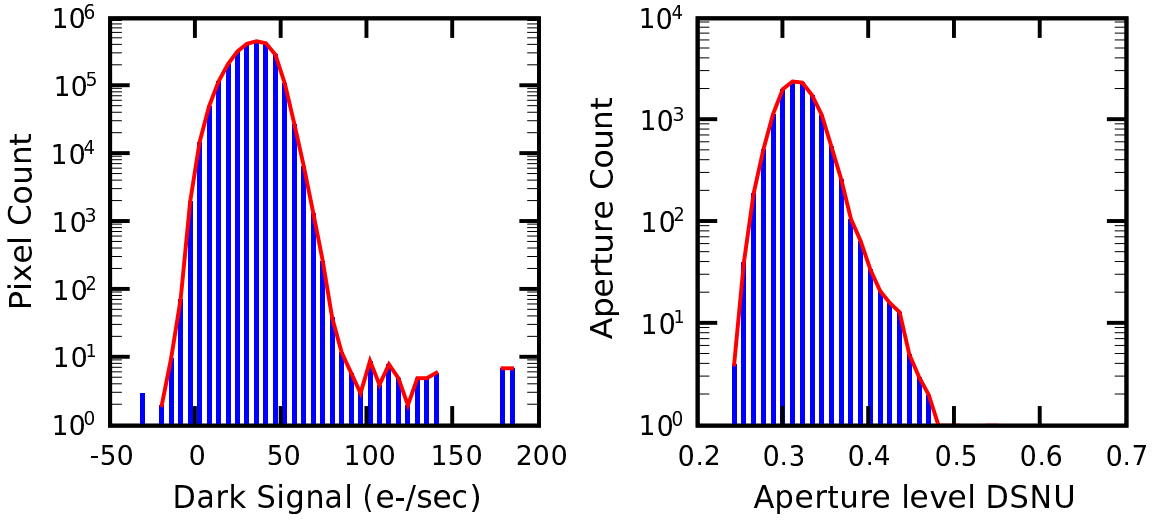



Figure 5.17: Photon transfer curve at 1/30 second exposure showing read noise of 5e- and PRNU limitation at high illumination.

Table 5.1: Summary of chip-level performance.

| **Parameter** | **Value** |
|---|---|
| Aperture count | 166×76 |
| Aperture format | 16×16 |
| Die size | 3.0×2.9 $mm^2$ |
| Maximum frame rate | 15 fps |
| ADC resolution | 10b |
| ADC FPN | 0.044 LSB (43$\mu$V) |
| ADC temporal noise | 0.26 LSB (254$\mu$V) |
| ADC INL/DNL | (0.46,-0.60)/(0.39,-0.57) |
| chip power | 10.45 mW |




(a) (b)

Figure 5.18: (a) Dark current histogram for all pixels in the image sensor, and (b) histogram for aperture-level DSNU.

Table 5.2: Summary of pixel-level performance at room temperature.

| **Parameter** | **Value** |
|---|---|
| Pixel size | 0.7$\mu$m |
| Well capacity | 3500e- |
| Conversion Gain | 165$\mu$V/e- |
| Responsivity at 550nm | 930e-/lux-s |
| QE at 550nm | 48% |
| CTE | 99.9% |
| Read Noise (RMS) | 5e- |
| Dark Signal at RT | 33e-/sec |
| DSNU (sigma/mean) | 0.35 |
| PRNU at sat. (sigma/mean) | 0.02 |
| Peak SNR | 35 dB |
| Dynamic Range | 57 dB |

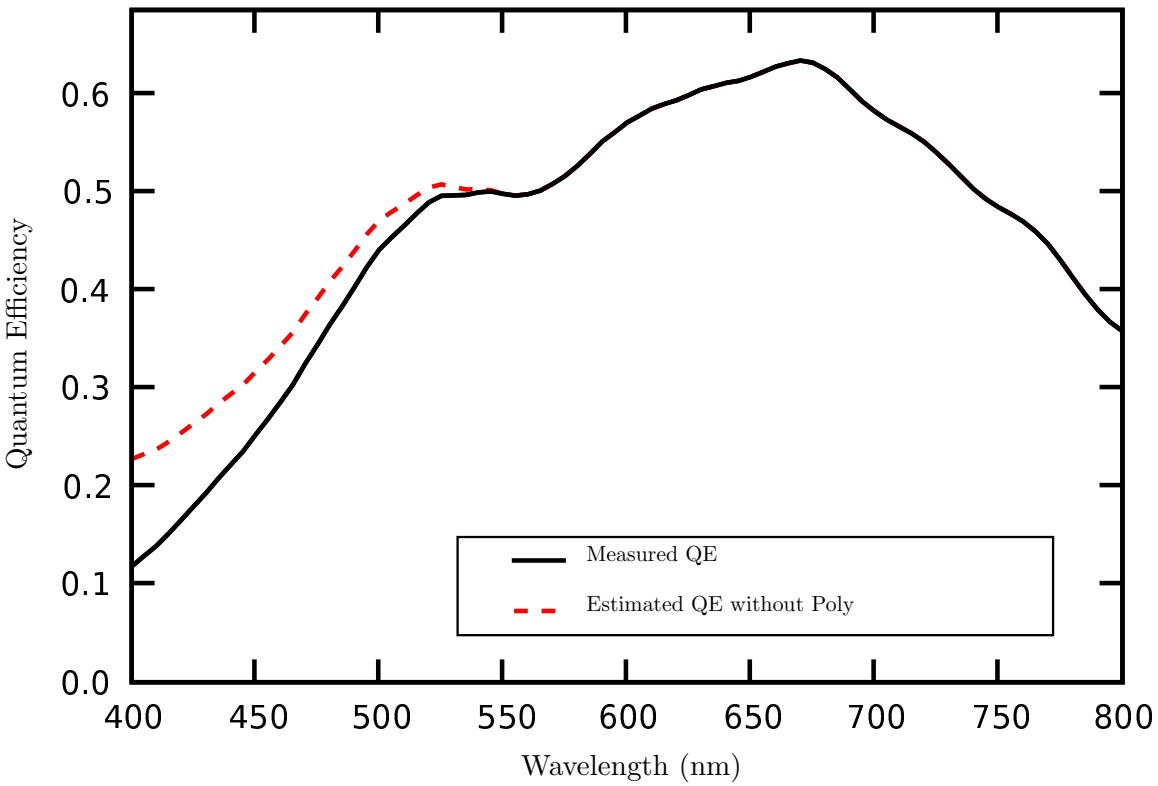
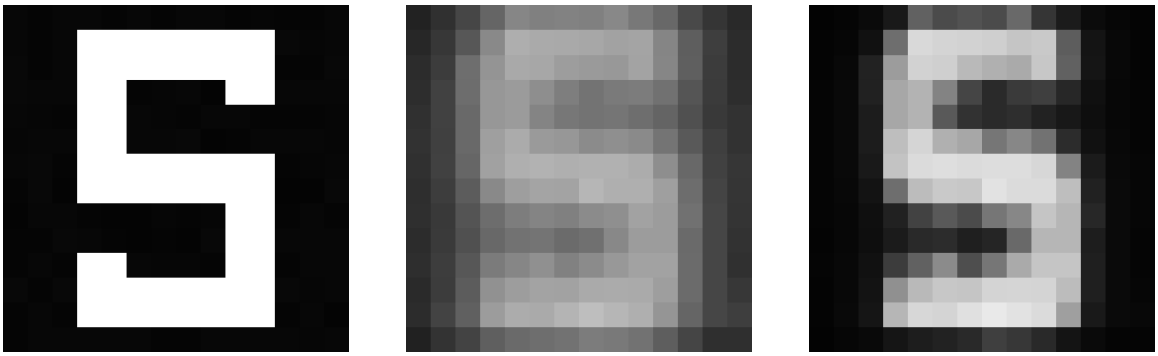



Figure 5.19: Measured QE at 5nm intervals showing strongest response at 650nm. The response in the blue region is reasonable due to the 65% polysilicon coverage of the active pixel area. The dashed line is an estimate of the QE if there were no absorption in the electrodes.

Figure 5.20: Sample image acquired with F/2.4 lens. The black grid is inserted by software to show the occluded area.



(a) (b) (c)

Figure 5.21: Sample images from a single 16×16 subarray. (a) Electrical image from fill-and-spill circuit. (b) Optical image projected from an LCD screen. (c) Optical image processed with added contrast enhancement.

the array. This demonstrates the charge confinement at pixel pitch. The optical image in Figure 5.21b demonstrates the imaging performance which is limited by the aperture of the camera lens at F/2.4. The spot size limit of this lens is between 2-3$\mu$m. However, with added contrast in Figure 5.21c we can clearly see the pattern.

To demonstrate the feasibility of color image reconstruction, subimages as they would appear through local optics are projected onto a subarray from an LCD display. The projected images are sorted by color to emulate the filter pattern shown in Figure 5.22. With sufficient overlap between subimages, the pattern is uniform for all color channels and empty space is filled in when apertures are projected to the focal plane. The raw data is obtained by cycling through all of the subimages for a given scene. Through simple reconstruction we produce the image in Figure 5.23. A global parameter is used to shift each subarray and to reconstruct the image by summing the overlapping views. This image shows that the pixels in the subarrays are adequate for image reconstruction and that excellent color separation between apertures can be achieved.

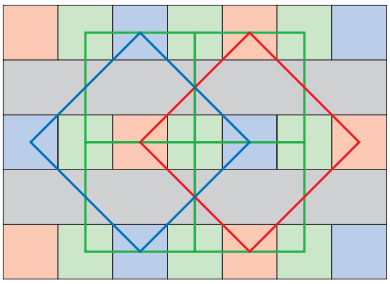


Figure 5.22: Color pattern for multi-aperture image sensor.

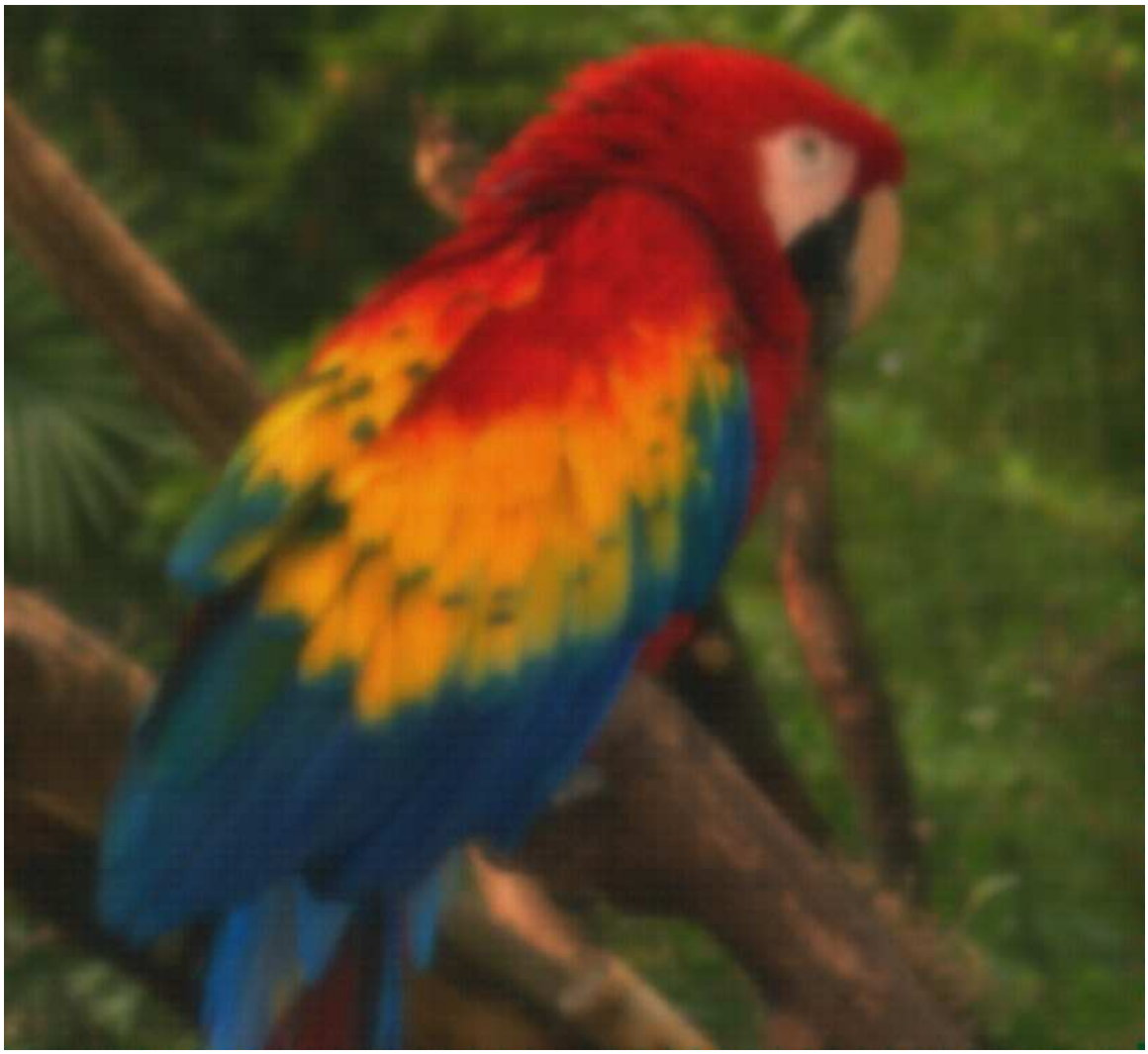
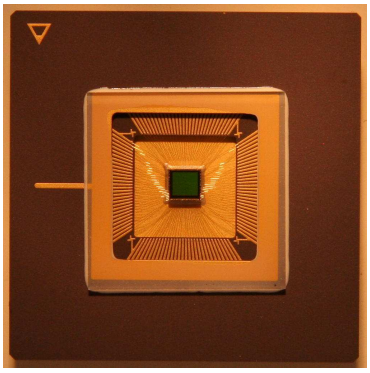
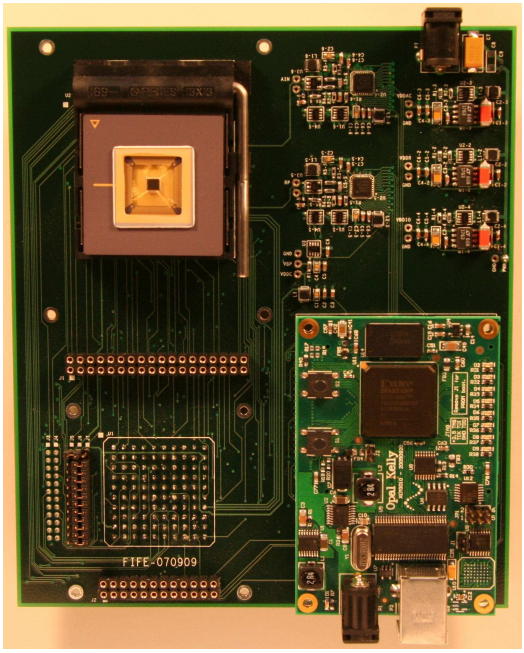


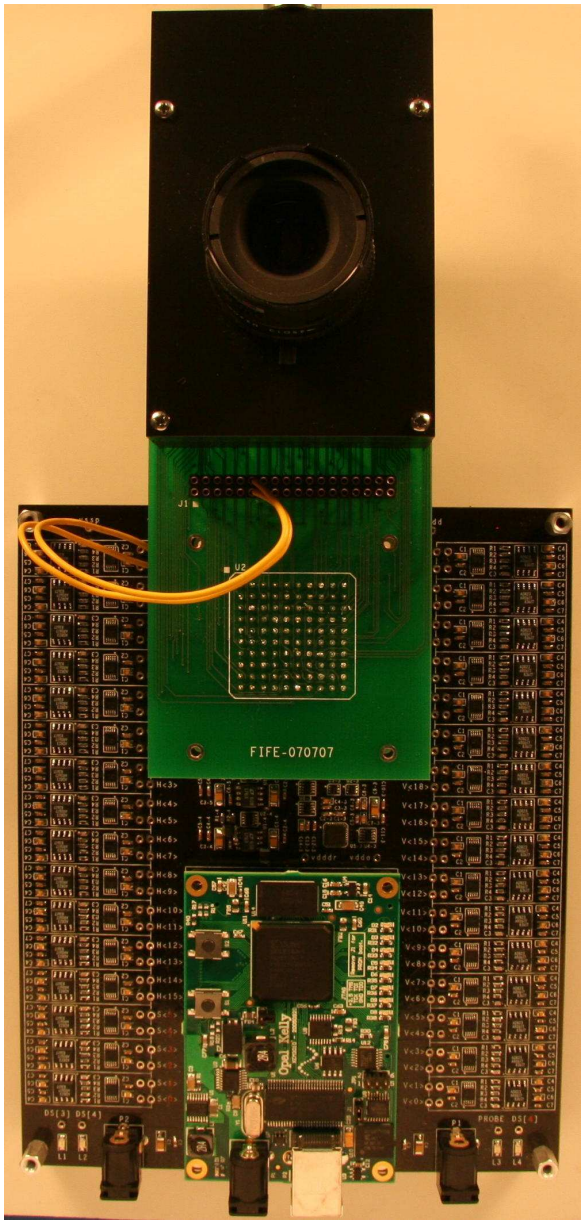
Figure 5.23: Color image captured by chip using projected subimages from LCD.



(a)




(b)




(c)

Figure 5.24: (a) Packaged chip, (b) adc test board and (c) camera for image capture.

# Chapter 6

# Conclusion

Design and characterization results of the first integrated multi-aperture image sensor have been reported. The results show good imaging performance with 0.7$\mu$m pixels using an FT-CCD subarray designed in deep submicron CMOS technology. Small pixel size is achieved by using ripple charge transfer, which allows for charge confinement at each electrode. The results suggest that further pixel size scaling is possible while maintaining acceptable performance. Improvements in pixel performance are expected with further modifications to the CMOS process to reduce dark current, improve charge transfer efficiency, and decrease PRNU. A variety of pixels including surface-channel, buried-channel, and pinned phase buried-channel have been implemented down to 0.5$\mu$m pitch to demonstrate the limitations and tradeoffs in designing deeply scaled pixels.

The multi-aperture architecture developed in this research creates a path for continued scaling of image sensor pixel count within the limits of practical optical formats, which may be needed to meet the demands of future imaging platforms. This is achieved in several ways:

- *Hierarchical readout*: Scaling pixel count presents a major challenge to readout speed and signal fidelity. A similar problem is solved in digital memories by breaking the memory array into smaller blocks and using hierarchical readout. The multi-aperture architecture applies this approach to image sensors. The gaps between the subarrays resulting from this hierarchical architecture is compensated for by using local optics with large enough magnification.

- *CCD subarray design*: CMOS pixels are becoming increasingly difficult to scale mainly due to optical considerations. As such, we proposed using an FT-CCD subarray design

with no metal occlusions in the pixel. Pixel size is minimized by using ripple charge transfer charge to allow for confinement between electrodes.

- *Per-Aperture CFA*: As discussed in Section 3.1.2, color cross-talk represents a major impediment to CMOS pixel scaling. The multi-aperture architecture solves this problem by using CCD pixel subarrays with no metal occlusions and a per-aperture color filter array.

- *Redundancy*: As pixel counts scale, it becomes increasingly difficult to achieve acceptable manufacturing yield without employing mechanisms for defect tolerance. This is achieved in the multi-aperture image sensor through the redundancy in the subimages.

In addition to enabling continued scaling of pixel count, the multi-aperture image sensor provides new important capabilities, including capturing depth information, relaxing the requirements on the objective lens, and imaging objects at close proximity without the need for objective optics. Furthermore, the architecture makes use of pixel sizes below the spot size limit of optics. It is generally understood that spatial resolution does not scale beyond approximately half the wavelength of the source of illumination. We have shown, however, that pixels made smaller than these dimensions still provide useful information, as demonstrated for the case of depth extraction. This work has taken the first step forward in building wavelength-sized pixels and demonstrating their use.

Realizing the benefits of the multi-aperture image sensor requires significant post-processing of the subimage data. This should not be the limiting factor to its adoption, however, because of the ever decreasing cost of computational resources.

# Appendix A

# Optimization for a Pinhole Camera

In order to create the smallest spot size possible in a lensless imaging system, the right trade off between geometry and diffraction must be made. Since the spot size can never be smaller than the aperture, it would be nice to make the aperture arbitrarily small. However, as the aperture gets smaller, there will be spreading due to diffraction. There is a clearly a trade off between geometry and diffraction. However, because of the spreading profile due to diffraction, it is not immediately obvious how to define the resolution of the system. There are various criteria that can be used, and since we are after some real numbers, we must decide on one in order to complete the optimization.

The *Sparrow resolution criterion* states that two equally strong incoherent point sources are just resolved when the dip in intensity between the two geometric images vanishes. In other words, the limit is when there is no longer two peaks.

The *Rayleigh Criterion* states that two incoherent point sources are just resolved when the Airy intensity pattern generated by one point source falls to zero at the peak of the second. Then it is a simple matter to calculate the minimum resolvable separation of two geometrical images produced by a diffraction-limited system with circular aperture by evaluating the Airy function.

Three similar methods are considered for the optimization of the pinhole camera shown in Figure A.1. Each method makes a trade off between diffraction and geometry. They differ slightly in the characterization of the diffraction and in the resolution criterion.

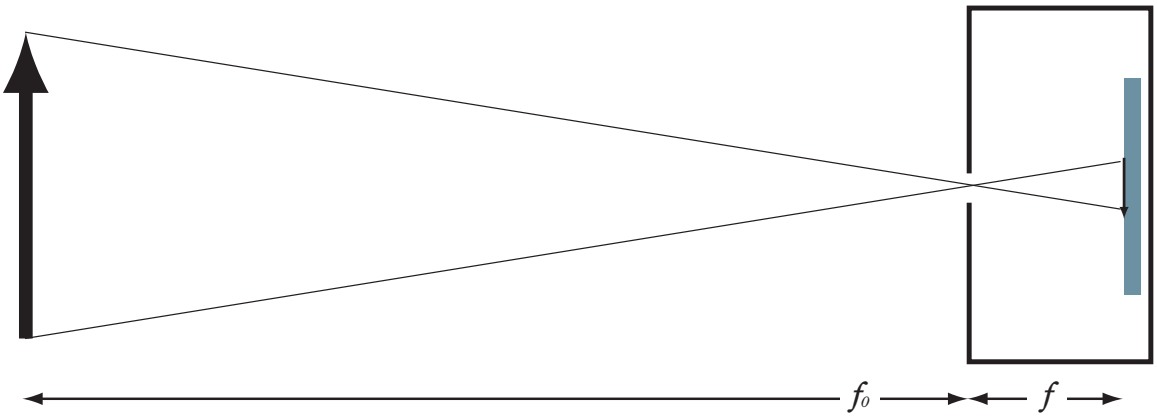



Figure A.1: Pinhole Camera.

## A.1 Method 1: Rayleigh Criterion with Fraunhofer Diffraction

Since the Rayleigh Criterion is easy to use, we can work with this in conjunction with a text book problem proposed by Goodman [51]. This method provides both insight into the trade off as well as a useful calculation.

### A.1.1 Geometric Consideration

We can start with the assumption that the pinhole is large enough to allow a geometric representation of the point spread function. Next, we can find the optical transfer function (OTF) and define the cutoff frequency of the camera to be the point where the first zero in the OTF occurs, according to the Rayleigh Criterion.

The point spread function is the projection of the pupil function onto the image plane. The Optical Transfer Function (OTF) is then the Fourier transform of the point spread function.

$$p(x,y) = A\text{circ}\left(\frac{r}{w}\right) \rightarrow s(u,v) = A\text{circ}\left(\frac{r'}{w}\right), \text{where } r' = \sqrt{u^2+v^2}.$$

The OTF is the Fourier transform of $s(u,v)$, where $\rho = \sqrt{f_x^2 + f_y^2}$.

$$\mathcal{F}(\rho) = 2\frac{J_1(2\pi\omega\rho)}{2\pi\omega\rho}$$

The first zero in the OTF occurs where $2\pi\omega\rho = 1.22\pi$, so the cutoff frequency is

$$\rho_{\text{cutoff1}} = \frac{0.61}{\omega}.$$

### A.1.2 Diffraction-limited Analysis

If we make the pinhole arbitrarily small, diffraction will dominate the spot size. We will use Fraunhofer diffraction with the assumption that the image is in the far field. The point spread function is now the scaled Fourier transform of aperture:

$$s(u,v) = A\left[2\frac{J_1(2\pi\omega r)/\lambda f}{2\pi\omega r/\lambda f}\right]^2$$

The OTF is the Fourier transform of $s(u,v)$, where $\rho = \sqrt{f_x^2 + f_y^2}$.

$$\mathcal{F}(\rho) = \frac{2}{\pi}\left[\arccos\left(\frac{\rho}{2\omega/\lambda f}\right) - \left(\frac{\rho}{2\omega/\lambda f}\right)\sqrt{1-\left(\frac{\rho}{2\omega/\lambda f}\right)^2}\right]$$

$$\rho_{\text{cutoff2}} = \frac{2\omega}{\lambda f}.$$

### A.1.3 Finding an Optimum

The geometric analysis indicates that decreasing the pinhole size increases the cutoff frequency, while the diffraction limited case shows that decreasing the pinhole size decreases the cutoff frequency. By combining both effects, we know that the geometric optics hold until the Fraunhofer diffraction is valid and starts to take over. The optimum will occur when both effects have equal impact on the frequency response. A reasonable approximation is to equate the two cutoff frequencies.

$$\begin{aligned}
\frac{0.61}{\omega} &= \frac{2\omega}{\lambda f} \\
\omega_{opt} &= \sqrt{(0.305)\lambda f} \\
d_{opt} &= (1.1)\sqrt{\lambda f}
\end{aligned}$$

Suppose that the focal length is $f = 10\mu m$ and that the wavelength of the source is $\lambda = 500nm$. This yields a $d_{opt} = 2.5\mu m$ which produces a minimum spacing size of $d_{space} = 4\mu m$.

## A.2 Method 2: Petzval and Mielenz

A more detailed and rigorous analysis of pinhole imaging with several criteria for resolution and a discussion of each is given in Mielenz [52]. The work makes reference to a "classical" paper by Petzval [53] in which he uses a fairly simple analysis to describe the progression of diffraction with geometry. First we can assume that $f_0 \gg f$ which approximates a plane wave at the aperture stop. See Figure A.2 for the progression of a point source through the aperture.

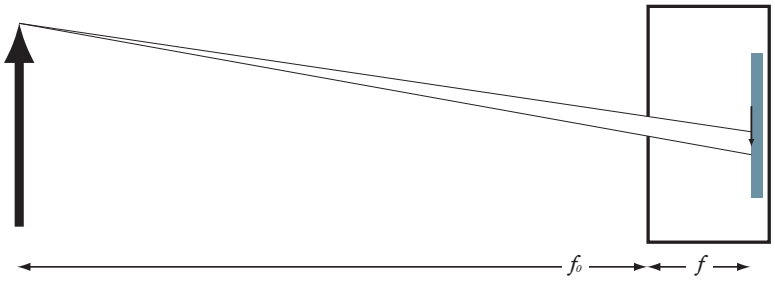



Figure A.2: Image of point source in pinhole camera.

We can assume that every point is spread by Fraunhofer diffraction which is roughly $\frac{2\lambda f}{a}$, where $a$ is the half width of the aperture. With this estimate, the assumption is that diffraction is dominate. In order to account for the geometry by estimation, we simply add in that portion. The total diameter of the spot size is then

$$d_{spot} = 2a + 2\lambda f/a.$$

The equation for $d_{spot}$ gives the distance from one zero to the next which estimates the width of the spot. By defining the minimum resolvable separation as the distance from the zero in the diffractive region to the start of the next geometric region, the minimum resolvable separation is

$$x_{min} = 2a + \lambda f/a.$$

Now there is a simple optimization as follows:

$$\begin{aligned}
\frac{\partial x}{\partial a} &= 2 - \lambda f/a^2 \rightarrow \frac{\partial x}{\partial a} = 0 \\
\lambda f/a^2 &= 2 \\
4a^2 &= 2\lambda f \\
2a &= \sqrt{2\lambda f} \\
d_{opt} &= (1.41)\sqrt{\lambda f}
\end{aligned}$$

This result is similar to the result in the previous analysis. Comparing the results with $f = 10\mu m$ and $\lambda = 500nm$ yields a $d_{opt} = 3.16\mu m$ which produces a minimum spacing size of $d_{space} = 6.32\mu m$.

## A.3 Method 3: Diffraction with Gaussian Beam

Both previous methods attempt to get a result with a simplification of the evolution of the diffraction pattern. A more accurate result, depending on the resolution criteria, can be obtained numerically by using the correct progression of diffraction such as that given by Fresnel diffraction in the near field. The criterion for resolution becomes more complicated because the evolution of the pattern itself changes with distance.

The diffraction that yields the smallest possible spot size is Gaussian beam diffraction. Working with the Gaussian is simple because the Fourier transform of a Gaussian is still Gaussian. Therefore, over the entire propagation of the beam, the wave remains Gaussian. We will approximate the system as a plane wave interacting with an aperture which produces a Gaussian beam with twice the $1/e$ half-width equal to the aperture size. See Figure A.3. The half-waist of the Gaussian propagates in $z$ as $w(z) = w_o\sqrt{1 + (z/z_R)^2}$, where $z_R = \pi w_o^2/\lambda$. The distance $z_R$ is referred to as the Rayleigh length. The radius of curvature starts out infinite like a plane wave and progresses towards a spherical wave with the point source at the origin. The radius of curvature is $R(z) = z + z_R^2/z$. It is straight forward to show that the optimum waist occurs at the Rayleigh length by taking the derivative of $w(z)$ with respect to $w_o$ and setting it equal to zero and solving for $z$. Then the optimum half-waist is found from the definition of the Rayleigh length.

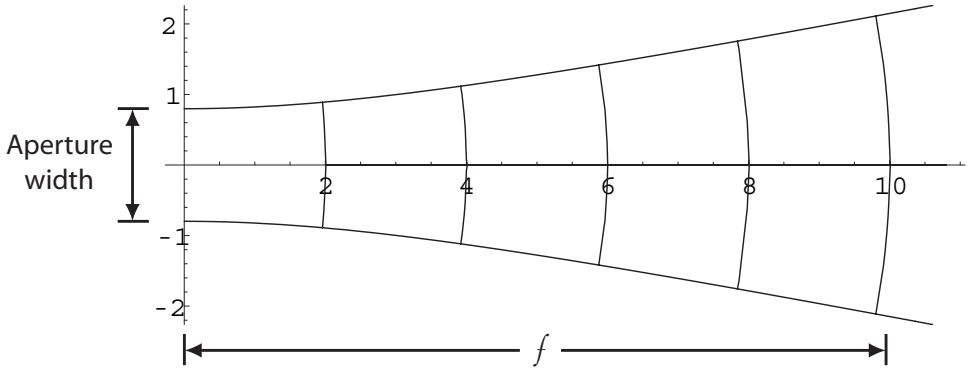



Figure A.3: Propagation of a Gaussian beam.

$$
\begin{aligned}
w_o &= \sqrt{\frac{\lambda}{\pi} z_R} \\
2w_o &= 2\sqrt{\frac{\lambda}{\pi} f} \\
d_{opt} &= 2\sqrt{\frac{\lambda}{\pi} f} \\
d_{opt} &= (1.128)\sqrt{\lambda f}
\end{aligned}
$$

The optimum aperture width is shown in Figure A.4. The result is similar to both of the previous cases. The optimization indicates that the aperture size is proportional to $\sqrt{\lambda f}$ in all three cases. With $f = 10\mu m$ and $\lambda = 500nm$, $d_{opt} = 2.52\mu m$ which produces a minimum spacing size of $d_{space} = 3.57\mu m$. The resolution criteria limit here occurs when two the Gaussian beam waists are just adjacent. This result is useful as a means of determining the absolute minimum spot size that could be produced from a pinhole aperture, even though this is not obtained in practice.

## A.4 Resolution as a Function of Distance to Source

The methods used to find the optimum aperture size for a pinhole camera made use of the assumption that the object was place at infinity. As the object becomes closer to the camera, the geometry changes as well as the optimization. If the lenless imaging system is

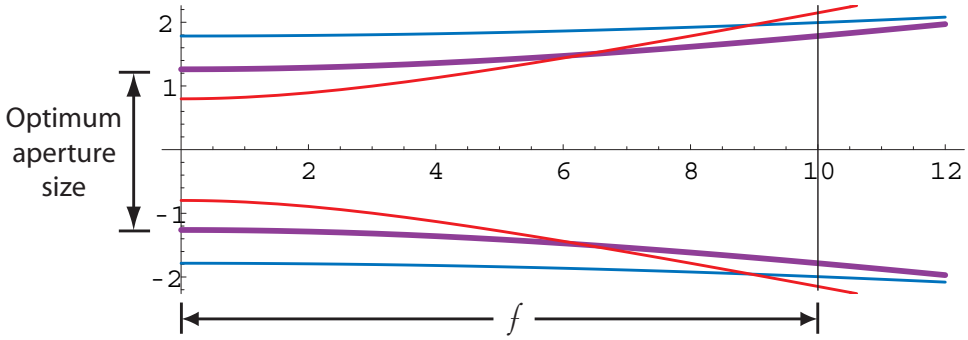



Figure A.4: Optimization for the initial beam waist of a Gaussian.

to operate at a fixed distance relatively close to the source, the difference can be significant. Mielenz [52] finds that the diffraction limit of such lenless imaging systems is given by $d_{opt} = \sqrt{2\lambda f f_0/(f_0 + f)}$. This result is essentially the same as that given in method 2 except that it accommodates a variable object distance. When the object is close to the camera, the minimum spot size is larger than it is when the object is imaged farther away. This is because the aperture must be stopped down to decrease the geometry of the beam which in turn increases its spread due to diffraction.

# Appendix B

# The Electrooptic Effect

## B.1 Controlling Refractive Indices with Electric Field

By applying an electric field to an optical medium, the refractive index may be altered for different polarization vectors. This example shows how the linear electooptic effect controls the refractive indices of a noncentrosymmetric crystal, Lithium Niobate $LiNbO_3$.

The electrooptic coefficients for $LiNbO_3$ are $r_{13} = 9.6$, $r_{22} = 6.8$, $r_{33} = 30.9$, $r_{42} = 32.6$ $(10^{-12} m/V)$. At $0.5\mu m$ illumination, $n_0 = 2.3410$ and $n_e = 2.2457$.

With the crystal oriented with the optical axis pointing out of the page along with the z-directed electric field, we can find the refracted angle $\theta_2$ as a function of $\theta_1$ and $E_z$. For this example, we assume a polarization in the $x$-$y$ plane. See Figure B.1.

If we take $l = 10\mu m$ and $\theta_1 = 30°$, we can get a good idea of how large the effect is within our scale of interest. For now, we will allow $Ez$ to vary from 0 to as much as $10^9 V/m$ (the breakdown field of a CMOS gate).

Lithium Niobate $LiNbO_3$, uniaxial crystal class $3m$.

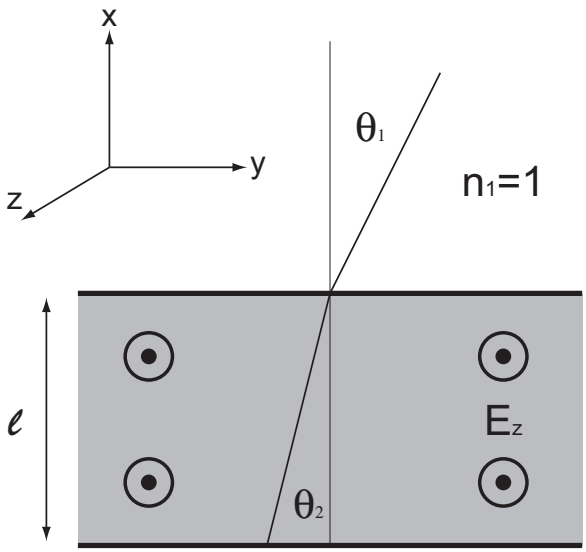



Figure B.1: $LiNbO_3$ slab with optical axis pointing out of page.

$$\begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \\ b_5 \\ b_6 \end{bmatrix} = \begin{bmatrix} 0 & r_{22} & r_{13} \\ 0 & r_{22} & r_{13} \\ 0 & 0 & r_{33} \\ 0 & r_{42} & 0 \\ r_{42} & 0 & 0 \\ r_{22} & 0 & 0 \end{bmatrix} \begin{bmatrix} E_x \\ E_y \\ E_z \end{bmatrix}$$

$$\begin{aligned} b_1 &= -r_{22}E_y + r_{13}E_z \\ b_2 &= r_{22}E_y + r_{13}E_z \\ b_3 &= r_{33}E_z \\ b_4 &= r_{42}E_y \\ b_5 &= r_{42}E_x \\ b_6 &= r_{22}E_x \end{aligned}$$

$$\Delta b_{ij} = \begin{bmatrix} (-r_{22}E_y + r_{13}E_z) & r_{22}E_x & r_{42}E_x \\ r_{22}E_x & (r_{22}E_y + r_{13}E_z) & r_{42}E_y \\ r_{42}E_x & r_{42}E_y & r_{33}E_z \end{bmatrix}$$

The impermiability tensor for any $E_x, E_y, E_z$ is:

$$\bar{\bar{b}} = \begin{bmatrix} b_0 + (-r_{22}E_y + r_{13}E_z) & r_{22}E_x & r_{42}E_x \\ r_{22}E_x & b_0 + (r_{22}E_y + r_{13}E_z) & r_{42}E_y \\ r_{42}E_x & r_{42}E_y & b_e + r_{33}E_z \end{bmatrix} \tag{B.1}$$

Assume a ray in $x$-$y$ plane with $E_z$ field and $E_x$, $E_y = 0$.

$$\bar{\bar{b}} = \begin{bmatrix} b_0 + r_{13}E_z & 0 & 0 \\ 0 & b_0 + r_{13}E_z & 0 \\ 0 & 0 & b_e + r_{33}E_z \end{bmatrix} \tag{B.2}$$

The index of refraction is found from the eigenvalues of the impermiability tensor.

$$n = \sqrt{1/\lambda} \quad \rightarrow \quad \lambda = 1/n_0^2 + r_{13}E_z \tag{B.3}$$

Considering only the index for polarization in the $x$-$y$ plane, $n_x = n_y$ so that the allowed index is:

$$\begin{aligned} n &= (1/n_0^2 + r_{13}E_z)^{-1/2} \\ &= n_0(1 + n_0^2 r_{13}E_z)^{-1/2} \end{aligned}$$

With the realistic limit of $n_0^2 r_{13}E_z \ll 1$,

$$n \approx n_0 - \frac{n_0^3 r_{13}E_z}{2} \tag{B.4}$$

The index varies linearly with $E_z$ as indicated by Equation B.4. Now we can use snell's law to find the deflection of the ray with a variation in $E_z$.

$$n_1 \sin\theta_1 = n_2 \sin\theta_2 \quad \rightarrow \quad \theta_2 = \arcsin\left(\frac{\sin\theta_1}{n_0 - \frac{n_0^3 r_{13} E_z}{2}}\right) \tag{B.5}$$

From Equation B.5, we find that $\theta_2$ increases as $E_z$ increases. With $E_z = 0$, $n_2 = 2.341$ and $\theta_2 = 12.33246°$. With $E_z = 10^9 V/m$, $n_2 = 2.2794$ and $\theta_2 = 12.6712°$. The change in the refractive index, $\Delta n$, is approximately 0.061 under the largest applied electric field. In this example, we created a deflection of about 0.3°. If the slab of $LiNbO_3$ were $10\mu m$ thick, the ray would walk about $62nm$.

# Bibliography


[1] Kwangho Yoon, Chanki Kim, Bumha Lee, and Doyoung Lee, "Single-chip CMOS image sensor for mobile applications," in *IEEE International Solid-State Circuits Conference, Digest of Technical Papers. ISSCC, Vol.1*, pp. 36–42, 2002.

[2] T. Hammadou, F. Boussaid, and M. Nilsson, "Low cost single chip CMOS camera for automotive application," in *Consumer Electronics, 2002. Digest of Technical Papers. International Conference on*, pp. 44– 45, 2002.

[3] I. Thompson-Bell, "Bare die imager," in *Electronics Manufacturing Technology Symposium, Twenty-Sixth IEEE/CPMT International*, pp. 27–32, 2000.

[4] G.Meynants, B.Dierickx, A.Alaerts, D.Uwaerts, S.Cos, D.Scheffer, and S.Noble (Eastman Kodak Company), "A 35mm 13.89 million pixel CMOS active pixel image sensor," in *IEEE Workshop on CCD and AIS*, 15-17 May 2003.

[5] B.J. Hosticka, W. Brockherde, A. Bussmann, T. Heimann, R. Jeremias, A. Kemna, C. Nitta, and O. Schrey, "CMOS imaging for automotive applications," in *Electron Devices, IEEE Transactions on*, pp. 173 – 183, January 2003.

[6] E.R. Fossum, "CMOS image sensors: electronic camera-on-a-chip," in *Electron Devices, IEEE Transactions on, Vol.44, Iss.10*, pp. 1689–1698, October 1997.

[7] Albert Theuwissen, *Solid-State Imaging with Charge-Coupled Devices.* Dordrecht: Kluwer, 1995.

[8] Junichi Nakamura *et al.*, *Image Sensors and Signal Processing for Digital Still Cameras.* Boca Raton, FL: CRC Press, 2006.

[9] S. Abdallah, B. Saleh, and A.K. Aboulsoud, "A general overview of solid state imaging sensor types," in *Photonics and Its Application at Egyptian Engineering Faculties and Institutes, 2002. Third Workshop on*, pp. 1–10.

[10] A.J.P. Theuwissen, "CCD or CMOS image sensors for consumer digital still photography?," in *VLSI Technology, Systems, and Applications, 2001. Proceedings of Technical Papers. 2001 International Symposium on*, pp. 168–171.

[11] B.S. Carlson, "Comparison of modern CCD and CMOS image sensor technologies and systems for low resolution imaging," in *Sensors, 2002. Proceedings of IEEE, Vol.1*, pp. 171– 176.

[12] L.J. Kozlowski and A. Luo, J.; Tomasini, "Performance limits in visible and infrared imager sensors," in *Electron Devices Meeting, 1999. IEDM Technical Digest. International*, pp. 867–870, December 1999.

[13] G. Agranov, V. Berezin, and R.H. Tsai, "Crosstalk and microlens study in a color CMOS image sensor," in *Electron Devices, IEEE Transactions on, Vol.50*, pp. 4– 11, January 2003.

[14] Ji Soo Lee, J. Shah, M.E. Jernigan, and R. Hornsey, "Characterization and deblurring of lateral crosstalk in CMOS image sensors," in *Electron Devices, IEEE Transactions on, Vol.50, Iss.12*, pp. 2361– 2368, December 2003.

[15] M. Furumiya, H. Ohkubo, Y. Muramatsu, S. Kurosawa, F. Okamoto, Y. Fujimoto, and Y. Nakashiba, "High-sensitivity and no-crosstalk pixel technology for embedded CMOS image sensor," in *Electron Devices, IEEE Transactions on, Vol.48, Iss.10*, pp. 2221–2227, October 2001.

[16] Sang Jong Park, Kwang Jin Kim, Joong Heon Kim, Seong Joon Lee, Sun Woong Woo, and Jeong Gun Lee, "Optical sensitivity and barrier property depending on ILD layers of CMOS image sensor device," in *VLSI and CAD, 1999. ICVC '99. 6th International Conference on*, pp. 211–213.

[17] H. Mutoh, "3-d optical and electrical simulation for CMOS image sensors," in *Electron Devices, IEEE Transactions on, Vol.50, Iss.1*, January 2003.

[18] Dun-Nian Yaung, Shou-Gwo Wuu, Ho-Ching Chien, Tzu-Hsuan Hsu, Chien-Hsien Tseng, Jeng-Shyan Lin, Jieh-Jang Chen, Chin-Hsin Lo, Chung-Yi Yu, Chia-Shiung

Tsai, and C.S. Wang, "Air-gap guard ring for pixel sensitivity and crosstalk improvement in deep sub-micron CMOS image sensor," in *Electron Devices Meeting, 2003. IEDM '03 Technical Digest*, pp. 16.5.1– 16.5.4, December 2003.

[19] C.-S.S. Lin, B.P. Mathur, and M.-C.F. Chang, "Analytical charge collection and MTF model for photodiode-based CMOS imagers," in *Electron Devices, IEEE Transactions on, Vol.49, Iss.5*, pp. 754–761, May 2002.

[20] E.-S. Eid, "Study of limitations on pixel size of very high resolution image sensors," in *Radio Science Conference, 2001. NRSC 2001. Proceedings of the Eighteenth National*, pp. 15–28, 2001.

[21] Hon-Sum Wong, "Technology and device scaling considerations for CMOS imagers," in *Electron Devices, IEEE Transactions on, Vol.43, Iss.12*, pp. 2131–2142, December 1996.

[22] Hon-Sum Philip Wong, "CMOS image sensors-recent advances and device scaling considerations," in *Electron Devices Meeting, 1997. Technical Digest*, pp. 201–204, December 1997.

[23] A. El Gamal, "Trends in CMOS image sensor technology and design," in *Electron Devices Meeting, 2002. IEDM '02*, pp. 805– 808.

[24] "Focus issue: Integrated computational imaging systems," in *Optics Express Vol. 11, No. 18*, September 2003.

[25] Joseph N. Mait, Ravi Athale, and Joseph van der Gracht, "Evolutionary paths in imaging and recent trends," in *Optics Express Vol. 11, No. 18*, pp. 2093–2101, September 2003.

[26] T.S. Huang, *Picture processing and digital filtering.* 1979.

[27] A.K. Jain, *Fundamentals of Digital Image Processing.* Englewood Cliffs: Prentice Hall, 1989.

[28] Wonpil Yu and Yunkoo Chung, "An embedded camera lens distortion correction method for mobile computing applications," in *Consumer Electronics, 2003. ICCE IEEE International Conference on*, pp. 400– 401, June 2003.

[29] T. Komatsu and T. Saito, "A high-resolution image acquisition method with defect-pixel recovery for solid-state image sensors," in *Image Processing, 2001. Proceedings*, pp. 1053–1056, October 2001.

[30] L. Landweber, "An iteration formula for fredholm integral equations of the first kind," in *Am.J.Math,73*, pp. 615–624, 1951.

[31] W. T. Cathey and E. Dowski, "A new paradigm for imaging systems," in *Applied Optics 41*, p. 6080, 2002.

[32] Kenneth Kubala, Edward Dowski, and W. Thomas Cathey, "Reducing complexity in computational imaging systems," in *Optics Express Vol. 11, No. 18*, pp. 2102–2107, September 2003.

[33] E.R. Dowski and W.T. Cathey, "Extended depth of field through wave-front coding," in *Applied Optics 34*, pp. 1859–1866, 1995.

[34] Jun Tanida, Tomoya Kumagai, Kenji Yamada, Shigehiro Miyatake, Kouichi Ishida, Takashi Morimoto, Noriyuki Kondou, Daisuke Miyazaki, and Yoshiki Ichioka, "Thin obervation module by bound optics (TOMBO): Concept and experimental verification," in *Applied Optics Vol. 40*, pp. 1806–1813, April 2001.

[35] K. Yamada, J. Tanida, Y. Kitamura, and Y. Ichioka, "An opto-electronic image capturing system using multiple-imaging CMOS sensor," in *Lasers and Electro-Optics, 2001. The 4th Pacific Rim Conference on, Vol.2*, pp. II–690– II–691.

[36] J Tanida, "Optoelectronic hybridization for compact imaging system," in *Lasers and Electro-Optics, 2003. CLEO/Pacific Rim 2003. The 5th Pacific Rim Conference on, Vol.2*, p. 424, December 2003.

[37] J. Tanida and K. Yamada, "TOMBO: thin observation module by bound optics," in *Lasers and Electro-Optics Society, 2002. The 15th Annual Meeting of the IEEE*, pp. 233– 234 vol.1.

[38] Jun Tanida, Rui Shogenji, Yoshiro Kitamura, Kenji Yamada, Masaru Miyamoto, and Shigehiro Miyatake, "Color imaging with an integrated compound imaging system," in *Optics Express Vol. 11, No. 18*, pp. 2109–2117, September 2003.

[39] H. Wei and T.D. Binnie, "High-resolution image reconstruction from multiple low-resolution images," in *Image Processing and Its Applications, 1999. Seventh International Conference on (Conf. Publ. No. 465), Vol.2*, pp. 596–600.

[40] N. George, "Lensless electronic imaging," in *Optics Communications, Jan 1, 1997 v133 i1*, p. p22(5).

[41] T. Massoud and S. Gambhir, "Molecular imaging in living subjects: Seeing fundamental biological processes in a new light," *Genes & Dev*, pp. 545–580, 2003.

[42] E.H. Adelson and J.Y.A. Wang, "Single lens stereo with a plenoptic camera," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 99–106, February 1992.

[43] Russell Thompson, Daniel Larson, and Watt Webb, "Precise nanometer localization analysis for individual fluorescent probes," *Biophysical*, pp. 2775–2783, May 2002.

[44] Ahmet Yildiz *et al.*, "Fluorophore imaging with 1.5-nm localization," *Science*, pp. 2061–2065, June 2003.

[45] H. Rhodes *et al.*, "CMOS imager technology shrinks and image performance," *IEEE Workshop on Microelectronics and Electron Devices*, pp. 7–18, 2004.

[46] R.H. Walden *et al.*, "The buried-channel charge coupled device," *Bell Systems Technical Journal*, vol. 51, pp. 1635–1640, 1972.

[47] L.J.M. Esser, "Peristaltic charge-couple device: A new type of charge transfer device," *Electronics Letters*, vol. 8, pp. 620–621, 1972.

[48] James Janesick, "Open-pinned phase CCD technology," *Proc.SPIE 1159*, 1989.

[49] Keith Fife, Abbas El Gamal, and H.-S. P Wong, "A 0.5$\mu$m pixel frame-transfer CCD image sensor in 110nm CMOS," *IEDM*, pp. 1003–1006, December 2007.

[50] W. H. White, D. R. Lampe, F. C. Blaha, and I. A. Mack, "Characterization of surface channel CCD image arrays at low light levels," *IEEE Journal of Solid-State Circuits*, vol. SC-9, pp. 1–14, February 1974.

[51] Joseph W. Goodman, *Introduction to Fourier Optics*. San Francisco, CA: McGraw-Hill, 1996.

[52] Klaus D. Mielenz, "On the diffraction limit for lensless imaging," in *J. Res. Natl. Inst. Stand. Technol. 104*, pp. 479–485, September 1999.

[53] J. M. Petzval *Wiener Akad. Ber*, pp. 26, 33, 1857.