Standard CMOS Active Pixel Image Sensors for Multimedia Applications

Alex Dickinson, Bryan Ackland, El-Sayed Eid and David Inglis
AT&T Bell Laboratories
Holmdel, NJ 07733

Eric R. Fossum
Jet Propulsion Laboratory, California Institute of Technology
Pasadena, CA 91109

Abstract

The task of image acquisition is completely dominated by CCD-based sensors fabricated on specialized process lines. These devices provide an essentially passive means of detecting photons and moving image data across chip. We argue that line widths in standard CMOS have been reduced to the point where it is practical to locate transistors - and hence provide gain - within each detector of an imager fabricated in CMOS. In this paper we describe the image sensors we have been building in this technology. They use far less power, permit random access and can be integrated with arbitrary digital CMOS logic to create low cost single chip video cameras with advanced functionality. Ongoing work includes the development of a color filter array process and the fabrication of a 1024x1024 pixel sensor for multimedia video and document capture applications.

1: Introduction

The use of video as a medium for personal communication is growing with the increasing availability of network bandwidth and video compression technology (e.g. ISDN and MPEG). The initial determinant of image quality is the video camera - where quality may be defined in terms of resolution, color rendition, and the "production values" enabled by advanced camera functionality (such as automatic head tracking).

Image sensors for use in electronic cameras are typically fabricated in highly specialized CCD processes. Being optimized for charge transfer these processes are not well suited to the integration of VLSI circuits with the image sensor to create true single chip camera systems. Attempts have been made to construct commercial-grade cameras with arrays of passive CMOS photo diodes [1] however these devices have high noise levels compared to CCD arrays, and scale poorly with array size.

In this paper we describe an approach that is radically different to both CCD and CMOS passive photo diode techniques: Active Pixel Sensors (APS). In the APS approach, each picture element (pixel) contains four (or three, depending on the pixel structure) transistors

in addition to the photo-sensing element (which may be either a photo diode or photogate). These devices provide gain within the pixel, greatly reducing noise, and allowing arrays to be scaled up with very little penalty. By fabricating the sensors in standard digital CMOS we are able to make use of aggressive scaling trends in these high volume processes. Pixels have been designed in a 2 micron process (40 micron pixel pitch), a 0.9 micron process (20 micron pitch) and a 0.5 micron process (10 micron pitch). CCD arrays are presently built on a 5 to 10 micron pitch. As line widths go through one more generation of reduction, the pixel pitch is expected to approach 5 microns, the expected limit for imager pixel size (independent of fabrication technology) due to optical diffraction limitations in the camera lens system.

The standard CMOS construction and SRAM-like architecture of the APS sensors give rise to the following advantages:

• Simple integration with other CMOS circuits (A/D converters, digital signal processors, digital bus interfaces, etc.) that allow for the reduction of overall system complexity and component count.
• Low power (CCD imagers require high voltage and high current clocks).
• Improved scaling capability: in a CCD, charge must be transferred serially through a large number of CCD elements, causing degradation of the signal with increasing array size.
• Electronic pan/zoom capability through the use of random access into the sensor area.
• Differential operation, in which the imager itself acts as a frame buffer for the previous frame, allowing only the difference between frames to be output for video compression, motion detection and image stabilization applications.

This combination of features suggests that the CMOS APS technology is well suited to the construction of cameras for multimedia applications where camera system cost, power, size and functionality are critical.

In this paper we will describe the design and testing of a 144x176 (QCIF) array and a 256x256 array.
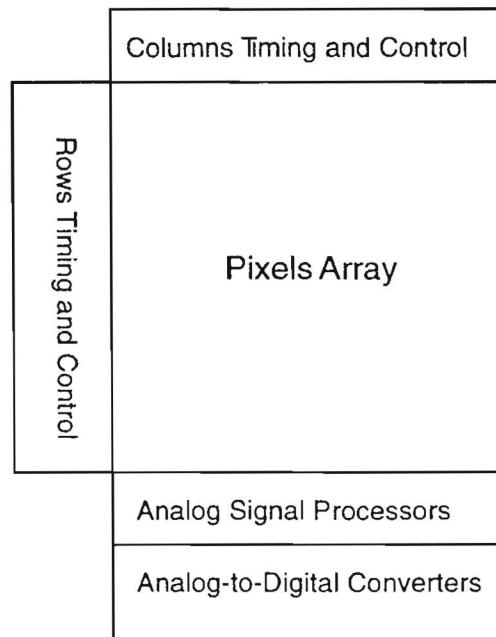
2: Active pixel sensor construction

2.1: Architecture

The general architecture of a CMOS active pixel image sensor, which has four main sections, is shown in Figure 1. The heart of the image sensor is an array of pixels. Within each pixel there are a photo-sensitive area and three or four transistors to buffer the photo-signal and enhance sensitivity and noise rejection. The size of the pixel array determines the spatial resolution of the image sensor. For example, a television format image sensor has a pixel array size of 640x480.

A CMOS active pixel array is randomly accessible in a similar manner to that of a random-access memory (RAM). The accessibility of the pixel array is achieved via two decoders: a row (horizontal) decoder and a column (vertical) decoder. The row decoder is to address one row at a time while the column decoder is to address one column at a time. The column decoder sequentially addresses all the columns during the period their row is being addressed by the row decoder. Therefore, the row decoder is the slow scan one and the column decoder is the fast scan one. The two decoders represent the minimum required

on-chip control and timing circuitry. Off-chip circuits can be used to step through all possible addresses and "raster-scan" the output of the image sensor.

The photo-signal is processed by an on-chip analog signal chain of circuits. There is an analog signal processor for each column. The primary function of the on-chip analog circuitry is to perform correlated double-sampling (CDS), thus eliminating reset noise and suppressing 1/f noise. Other functions, such as suppressing fixed pattern noise, can be

Columns Timing and Control	
Rows Timing and Control	Pixels Array
	Analog Signal Processors
	Analog-to-Digital Converters

**Figure 1: Architecture of a CMOS Active Pixel Image Sensor.**

added as well. The output of the analog circuitry is a pair of analog signals. The first analog signal represents the photo-signal level while the second one represents the dark signal level. The analog video signal is the difference between these two signals.

Because of the use of standard CMOS construction, it is possible to directly integrate an analog-to-digital converter (ADC) with the camera array. Depending on the real-estate and power requirements, a column-parallel approach or a serial approach can be adopted. In the column-parallel approach, an ADC is devoted to each column of the pixel array, which fits nicely with the column-wise analog signal processor architecture. On the other hand, in the serial approach, a single ADC is integrated at the output of the analog signal chain circuitry.

2.2: Pixel design and operation

The double-polysilicon CMOS active pixel design is shown in Figure 2 [2]. The double-polysilicon structure provides the required electrical coupling between the semiconductor substrate surface regions under the photo-gate (PG) and the transfer gate (TG). The substrate surface potential under PG is controlled by the voltage applied to PG. Likewise, the substrate surface potential under TG is controlled by the voltage applied to TG. If the gap between PG and TG is not narrow enough, the substrate surface potential under this gap will constitute a barrier (electrical potential barrier) for the transfer of the charge signal from under PG to under TG, resulting in the failure of the signal read-out. The double-

polysilicon structure permits the overlapping of PG and TG as each is on a different polysilicon level. Consequently, the gap between PG and TG is as narrow as the thickness of the silicon dioxide layer between them, which is in the order of a few tens of nano-meters. This very narrow gap and the coupling capacitance due to overlapping provide the required electrical coupling.
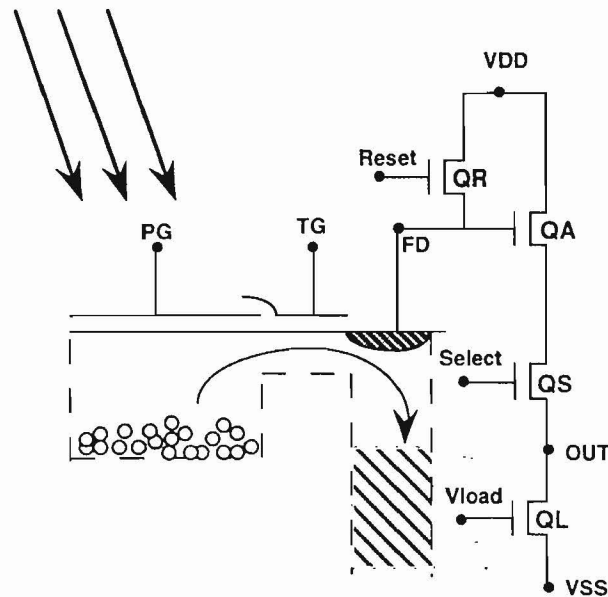
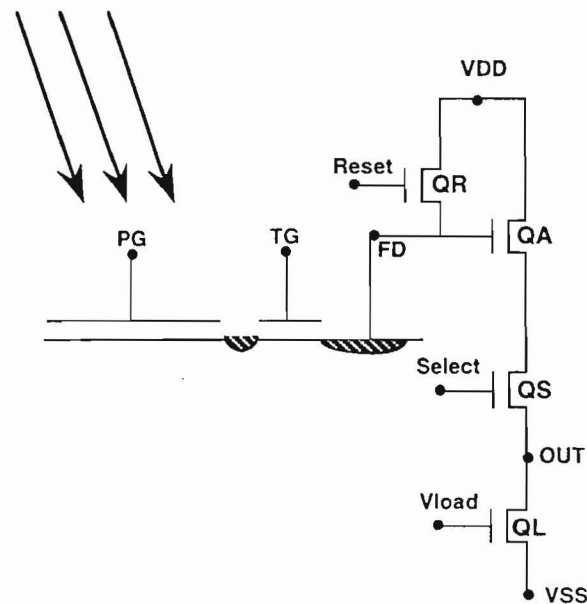

Figure 2: Double-Polysilicon Active Pixel.

The operation of the double-polysilicon pixel has two phases. In the first phase, integration, the generated photo-charge carriers are collected under the photo-gate (PG) for a predetermined period (integration time). This is done by clocking PG to a high voltage level (approximately VDD). In this phase, the transfer gate, TG, is turned off. In the second phase of operation, read-out, QR is pulsed on and off. This causes the potential of FD to float at a level approximately equal to VDD less the threshold voltage. Then, the bias of PG is changed to approximately VSS, causing the transfer of the charge signal into FD. This charge signal transfer causes the potential of FD to deviate from its approximately VDD value (reset level) to another value (signal level). This potential deviation (the difference between the reset and signal levels) is proportional to the incident light intensity and constitutes the video signal. A source-follower consisting of an active transistor (QA1) and a load transistor (QL1) is used to buffer the pixel FD node from the output node (OUT1) which is common (along with QL1) to a column of pixels within the image sensor pixel array. A select transistor (QSR) is used to select the pixel (along with its common row of pixels) for read-out. In this phase, the transfer gate, TG, is turned on to allow the transfer of the signal charge from under PG into FD. However, it is turned off right after the completion of this transfer and before PG is clocked high for the following integration time, ensuring that none of the signal charge will transfer back to under PG causing image lag.

The transfer gate, TG, should be clocked on and off as described above for the optimum operation of the pixel. However, if TG is dc biased such that it is slightly conducting, the pixel operates satisfactorily. DC biasing the transfer gate makes the operation simpler by eliminating the need for one control clock and the associated driving circuitry, but it may

cause some of the signal charge to transfer back to under PG, resulting in image lag. This effect is more pronounced at high signal levels.

If the above structure (shown in Figure 2) were to be fabricated using only one level of polysilicon, the gap between PG and TG would be in the order of a micron, too wide to provide the required electrical coupling. An active pixel design which employs only one level of polysilicon and provides the required electrical coupling is shown in Figure 3. In this design, a transfer transistor (QT) replaces the transfer gate (TG). This is equivalent to introducing a diffusion region (called coupling diffusion) between PG and TG of the structure outlined in the previous section. This diffusion region functions as a conducting channel between the substrate surface under PG and the channel of the transfer transistor QT, thus providing the required electrical coupling.

The introduction of the coupling diffusion has the effect of increasing the kTC noise of the pixel. This effect can be minimized by making the coupling diffusion capacitance as low as possible. This coupling diffusion capacitance can be made as low as few to several femto-Farads. This is equivalent to kTC noise in the order of few tens of electrons.

Figure 3: Single-Polysilicon Active Pixel.

The single-polysilicon CMOS active pixel can be operated in a fashion similar to that of its double-polysilicon counterpart. For example, the transfer transistor, QT, can be turned on and off in a similar manner to that of the transfer gate, TG, for the optimum operation of the pixel. On the other hand, the transfer transistor can be dc biased such that it is slightly conducting, resulting in the trade-off between operation simplicity and image lag effect as described above.

2.3: On-chip timing and control circuitry

The most straight forward way to operate an active pixel image sensor is the monolithic integration of two shift registers; a row (horizontal) shift register and a column (vertical) shift register. The row shift register, which is the slow scan one, selects one row at a time. On the other hand, the column shift register, which is the fast scan one, selects one column

at a time. The column shift register sequentially addresses all the columns during the period their row is being addressed by the row shift register. This method of operating an active pixel image sensor is attractive because of its simplicity. However, it lacks the very significant functionality of random accessibility, which may not be critical to consumer products.

Alternatively, an active pixel image sensor array may be randomly accessible in a similar fashion to that of a random-access memory (RAM). The accessibility of the pixel array is achieved via two decoders; a row (horizontal) decoder and a column (vertical) decoder. The slow scan row decoder is to address one row at a time while the fast scan column decoder is to address one column at a time. The column decoder sequentially addresses all the columns during the period their row is being addressed by the row decoder. Off-chip circuits can be used to progressively step through all possible addresses and "raster-scan" the output of the image sensor. However, two on-chip counters can be employed to interface the two decoders. Counters with load and clear control signals add the very significant functionality of random accessibility. In this mode of operation, instead of reading-out the whole image frame, a random window (sub-frame) can be read-out. The starting horizontal and vertical addresses of the window can be loaded into the counters in conjunction with the two load control signals. Moreover, an analog addressing technique can be utilized. In this technique, these two addresses are loaded in an analog form to two on-chip analog-to-digital converters (ADC). The digital output of each of these two ADCs is in turn fed into the corresponding counter. The analog addressing technique has the advantage of requiring only two input pads instead of, for example, 16 input pads (for the case of a 256x256 pixel array). On the other hand, the horizontal and vertical dimensions of the window can be determined via the timing of the two clear control signals.

Finally, clock generation as well as digital programmability circuitry can be monolithically integrated. Thus, the final chip is a camera rather than an image sensor. Read-out rate, integration time, and windowing functions are to be downloaded to the chip as part of mode control in an initial set-up phase of chip operation. Once set, the chip operates in the commanded mode until further programming is received.

2.4: On-chip analog circuitry

The photo-signal is processed by an on-chip analog signal chain of circuits shown in Figure 4 [3]. There is an analog signal processor for each column. The primary function of the on-chip analog circuitry is to buffer the pixel read-out node (FD) and perform correlated double-sampling (CDS). The first stage of this analog signal processor is the in-pixel source-follower. The in-pixel source-follower, which consists of an active transistor (QA) and a load transistor (QL), is used to buffer the pixel read-out node (FD) from the output node (OUT) which is common (along with QL) to a column of pixels within the image sensor pixel array. A select transistor (QS) is used to select the pixel (along with its common row of pixels) for read-out. Thus, the pixel read-out node is local to its pixel and buffered from any other read-out nodes within the image sensor pixel array. Consequently, the capacitance of the read-out node is significantly reduced, resulting in significant improvements of read-out sensitivity and noise. For an 128x128 image sensor pixel array, the reduction in the capacitance of the read-out node is four orders of magnitude compared to that of a design where the read-out node is common for all pixels within the image sensor array, and two orders of magnitude compared to that of a design where the read-out node is common for all pixels within a row or a column of the image sensor array [3]. This reduction in read-out node capacitance is more for larger size image sensor arrays. A re-

duction of the read-out node capacitance results in an increase of the read-out sensitivity by the same amount and a reduction of the read-out kTC noise by the square root of this amount.

The second stage of the analog signal processor is a parallel pair of sample-and-hold circuits. The first sample-and-hold circuit handles the reset voltage level and consists of a sampling switch (QRS) and a holding capacitor (CRH). Similarly, the second sample-and-hold circuit handles the signal voltage level and consists of a sampling switch (QSS) and a holding capacitor (CSH). The third and final stage of the analog signal processor is a parallel pair of source-followers, each drives an output pad. The first source-follower handles the reset voltage level and consists of an active transistor (QA2) and a load transistor (QL2). Similarly, the second source-follower handles the signal voltage level and consists of an active transistor (QA3) and a load transistor (QL3). Two select transistors (QSCR and QSCS) are used to select the pixel, along with its common column of pixels.
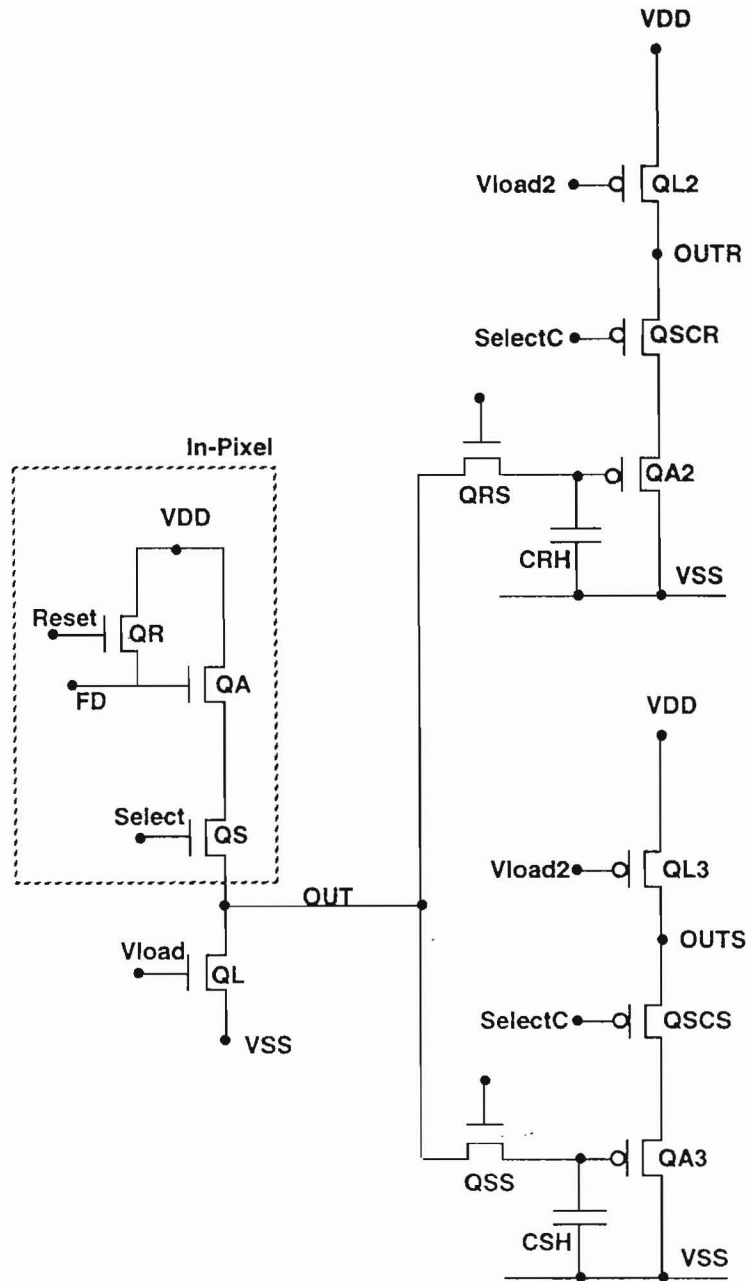
Figure 4: On-Chip Analog Signal Processor.

The output of the analog circuitry is a pair of analog signals. The first analog signal (OUTR) represents the dark signal level while the second one (OUTS) represents the photo-signal level. The analog video signal is the difference between these two signals. This difference operation (OUTR-OUTS) is done off-chip. This approach eliminates kTC noise through CDS, as well as fixed pattern noise (FPN) due to pixel transistor offsets, and suppresses 1/f noise. However, it introduces additional kTC noise due to the two hold capaci-

tors and FPN due to unmatched column source-follower transistors. This FPN can be suppressed via additional on-chip analog circuitry or off-chip signal subtraction techniques.

It should be noted that the physical layout of the analog signal processor shown in Figure 4 is at the bottom of each of the columns of the pixel array. The only exception is the in-pixel active transistor (QA), the row select transistor (QS), and the reset transistor (QR). Each pixel in the array has its local active transistor (QA), row select transistor (QS), and reset transistor (QR).

2.5: On-chip analog-to-digital conversion

The final main section of the CMOS active pixel image sensor is the on-chip analog-to-digital converter (ADC). Depending on the real-estate and power requirements, a column-parallel approach or a serial approach can be adopted. In the column-parallel approach, an ADC is devoted to each column of the pixel array, which fits nicely with the column-parallel analog signal processor architecture. On the other hand, in the serial approach, a single ADC is integrated at the output of the analog signal chain circuitry. We have found that the trade space for lower resolution (8-bit) ADC architectures to be somewhat flat so that several approaches for on-chip A/D conversion are being explored. A column-parallel approach is being used to reduce bandwidth and total power requirements. An image sensor with column-parallel single-slope ADCs for 8-bit resolution has been designed and fabricated, and is currently under test.

2.6: Differential mode of operation

To operate the device in differential mode, in which the difference between sequential frames appears at its output, we make use of the fact that at the end of a normal cycle, the floating diffusion output node acts as a dynamic storage node onto which has been placed a charge representing the optical input of the photogate. In effect the array of storage nodes is an implicit analog frame buffer, storing the entire frame. By altering only the timing, it is possible to generate the difference signal. Rather than starting with a reset operation, the value on the output node (the "old" frame) is read out to one of the sample-and-hold capacitors. The reset is then performed, followed by a normal read operation (the "new" frame). This signal value is then stored on the other sample-and-hold capacitor. The two capacitors are then fed to a differential amplifier in the normal manner, however the output of that amplifier is now the frame-to-frame difference. Note that the hardware required for both normal and differential modes is identical, only signal timing is altered.

Although there is some charge leakage from the storage node during the integration period, under normal exposure conditions this does not have a significant effect on differential image quality. In a motionless scene, the differential output signal is 0V. Any movement in the scene causes a non-zero output to appear. By correlating the non-zero output with the address currently being applied to the row and column decoders, it is possible to determine the position of the motion in the scene. This output may be used as the basis for compression, motion detection, and image stabilization operations without the use of an explicit frame buffer or ALU.

3: Test devices

3.1: A QCIF (176x144) image sensor

The first APS device designed for the AT&T 0.9μm CMOS process was a QCIF (144 x 176) image sensor with a 20 μm x 20 μm pixel size. Several different pixel layouts, including a single-poly pixel and a double-poly pixel, were deployed in an array configuration. The resultant pixel fill-factor is approximately 25%. Two 8-input NAND decoders, with a 20 μm pitch, were used for row and column selection with external circuits used to step through all possible addresses and "raster-scan" the output of the image sensor. The analog signal processor, which was laid out at the bottom of the pixel array, has a 20 μm pitch as well. No fixed pattern noise suppression circuitry was added to the analog signal processor for this particular set of image sensors. The output of the chip is a pair of analog signals. The first analog signal represents the photo-signal level while the second one represents the dark signal level. The analog video signal is the difference between these two signals. This difference operation was done off-chip. The total size of the 44-I/O pad chip is approximately 4.5 mm x 5.0 mm.

The CMOS active pixel image sensor was successfully demonstrated. It was operated at two different power supply voltage levels; 5.0 V and 2.5 V. At 5.0 V, the single-poly image sensor exhibited a video signal saturation level of approximately 1500 mV. The conversion gain (read-out sensitivity) was measured to be approximately 7 μV per electron. This corresponds to a read-out node capacitance of approximately 23 fF, which is in a very good agreement with the value extracted from physical design. The measured voltage signal saturation level and conversion gain yields a charge signal saturation level of approximately 215 thousand electrons per pixel. However, the pixel charge capacity is approximately 6 million electrons. This means that the signal saturation level is limited by the output source-follower(s) rather than by the pixel charge capacity. The measured rms read-out noise is approximately 290 μV (equivalent to approximately 42 electrons), which is in a very good agreement with the extracted value of the read-out node capacitance. The dynamic range, which is defined as the ratio of signal saturation level to rms read-out noise level, is approximately 74 dB. This is equivalent to better than 12 bits. The peak-to-peak fixed pattern noise was measured to be approximately 25 mV, or approximately 1.7% of saturation level. The dark signal (thermally generated signal) was measured to be approximately 160 mV per second at room temperature. At a pixel rate of approximately 320 thousand pixels per second (13 frames per second), the measured power dissipated in the chip is approximately 68.5 mW.

At 2.5 V operation, the single-poly image sensor exhibited a video signal saturation level of approximately 225 mV. The conversion gain was measured to be approximately 5 μV per electron. This corresponds to a read-out node capacitance of approximately 32 fF. The measured rms read-out noise is approximately 84 μV (equivalent to approximately 17 electrons). The dynamic range is approximately 68 dB, which is equivalent to better than 11 bits. The peak-to-peak fixed pattern noise was measured to be approximately 10 mV, or approximately 4.4% of saturation level. The dark signal was measured to be approximately 22 mV per second at room temperature. At a pixel rate of approximately 320 thousand pixels per second, the measured power dissipated in the chip is approximately 7.5 mW.

3.2: A 256x256 image sensor

A 256x256 CMOS active pixel image sensor has been designed, fabricated, and tested. The need for a double-poly process was eliminated by using a single-poly pixel and a MOSFET holding capacitor (in lieu of a double-poly one). Pixel and peripheral circuitry designs similar to those used in the QCIF array were employed, yielding performance parameters very similar to those described above. However, the 256x256 image sensor has the

additional functionality of random-accessibility, thus allowing a simple implementation of electronic pan and zoom which was successfully demonstrated. This was achieved by the monolithic integration of two counters interfacing the two decoders. Each of the two counters has load and clear control signals. In this mode of operation, instead of reading-out the whole image frame, a random window (sub-frame) is read-out. The starting horizontal and vertical addresses of the window are loaded into the counters in conjunction with the two load control signals. The horizontal and vertical dimensions of the window are determined via the timing of the two clear control signals.

Figure 5 shows the quantum efficiency (QE) of the APS compared to that of two commercial CCD devices. The APS device fabricated in the standard single-poly digital CMOS process has comparable performance over most of the visible spectrum, however there is reduced response towards the blue end. Conversion gain within the pixel was measured at 6.75 uV per electron and the saturation level is 1.5V output-referred. The total size of the 49-I/O pad chip is approximately 6.2 mm x 6.4 mm.
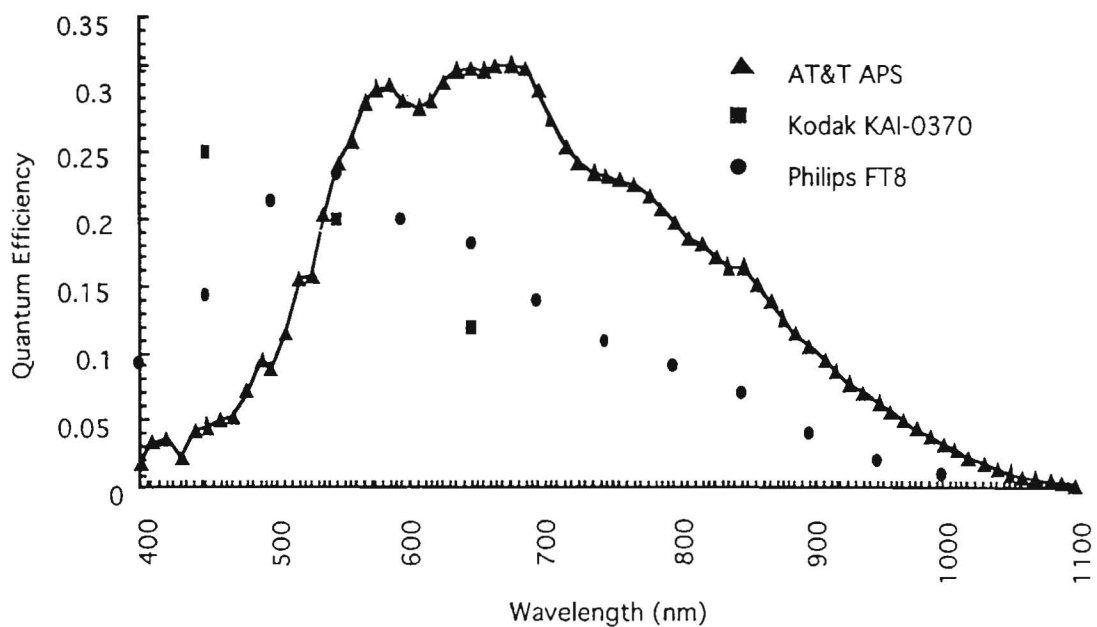

Figure 5: Quantum efficiency of the CMOS APS and two commercial CCD devices.

4: Conclusions

Our experience with the construction of CMOS APS devices suggests that they offer image quality similar to that of CCD cameras, without having to resort to a specialized silicon process for fabrication. This, together the with low power and high functionality of the APS technology make it ideal for multimedia camera construction. We have demonstrated these capabilities through the design and fabrication of QCIF and 256x256 format APS devices. Ongoing work includes the development of a color filter array process and on-chip A/D conversion - devices incorporating both these technologies are presently under test. Recently we began fabrication of a 1024x1024 APS sensor in 0.5µm digital CMOS. This sensor is expected to cost an order of magnitude less than an equivalent CCD device, and will be used in high resolution color video and document imaging multimedia applications.

References

[1] D. Renshaw, P. Denyer, G. Wang, and M. Lu, "ASIC Vision", 1990 IEEE CICC Proc., pp 7.3.1-7.3.4 (1990).

[2] E.R. Fossum, "Active Pixel Image Sensors -- Are CCDs Dinosaurs?", Proc. SPIE vol. 1900, pp 2-14 (1993).

[3] S. Mendis, S. Kemeny, and E.R. Fossum, "A 128 X 128 CMOS Active Pixel Image Sensor for Highly Integrated Imaging Systems", 1993 IEEE IEDM Tech. Dig., pp 583-586 (1993).