Optimization of Noise and Responsivity in CMOS Active Pixel Sensors for Detection of Ultra Low Light Levels

Orly Yadid-Pecht*, Barmak Mansoorian**, Eric R. Fossum** and Bedabrata Pain*

*Jet Propulsion Laboratory, California Institute of Technology
Pasadena, California, 91109

** Photobit LLC, La Cresenta, California 91214

Abstract

In this paper, we present results of the investigation of the design and operation of CMOS active pixel sensors for detection of ultra-low light levels. We present a detailed noise model of APS pixel and signal chain. Utilizing the noise model, we have developed APS pixel designs that can achieve ultra-low noise and high responsivity. We present results from two test chips, that indicate (1) that less than 5 electrons of read noise is possible with CMOS APS by reducing the size of the pixel transistors, and (2) that high responsivity can be achieved when the fill-factor of the photodiode is reduced.

1. Introduction

The CMOS active pixel sensor (APS) technology has recently attracted interest due to its low power consumption and its abilities to integrate the sensor control, readout electronics, and additional functionality on the imager chip[1-6], while retaining excellent imaging performance. The imager characteristics, such as quantum efficiency, read noise, fixed-pattern noise, imaging formats, pixel sizes etc. are acceptable for most commercial and industrial applications. However, the typical read noise in APS - in the range of 10-15e - is considered too high for some space science applications. Space science sensors are often required to operate at below 5 electrons read noise - scientists preferring sub-electron read noise in many space-telescope based environments. At the same time, scientists also desire ever increasing spatial resolution (pixel size less than 10 microns).

In this paper, we analyze the effect of pixel design parameters, specifically pixel fill factor and size of pixel amplifier, on APS read noise and responsivity. We present a detailed model of noise in APS, demonstrate operation of a small-pixel APS with less than 5 electrons read noise, and present both theoretical and test results that contradict the common belief that lower noise requires larger amplifier size. Another common belief is that the larger the fill-factor, the higher the sensor responsivity. However, responsivity is a complex function of photoelectron collection efficiency, detector capacitance, and quantum efficiency, and it is not immediately clear that the common belief is correct. Here we present test results contrary to the popular belief and demonstrate that high responsivity is indeed achieved with a small fill-factor, especially for photodiode type APS.

2. Noise in APS

The basic APS pixel, with a partial representation of the signal chain, is shown in Figure 1. A detailed description of the APS readout chain is given elsewhere[1]. The pixel contains a photogate (PG) with a floating diffusion (FD) output separated by a transfer gate (TX). In the photodiode (PD) type pixel, the TX transistor is omitted and the floating diffusion is actually extended to be the photon collecting area. The pixel also contains a reset transistor, the input FET (M1) of the source follower (SF), and

a row selection transistor. The readout circuit, common to an entire column, includes the load transistor (M2) of the SF, and two sample-and-hold circuits. Each sample-and-hold circuit contains a MOSFET switch (MSHS, MSHR), a capacitor, a column selection transistor, and a SF input transistor (Mp1), where the corresponding load transistor (Mp2) is common to all columns. The readout operation consists of sampling the voltage at the output of the pixel source follower twice - once following the pixel reset, and then following the transfer of charge from the photogate. The sampled voltages are stored on the column capacitors and are sequentially read out with the p-channel column source followers.

There are several sources of noise in the readout signal chain - both temporal and spatial, the latter also known as fixed pattern noise (FPN). FPN is caused by MOSFET threshold mismatches in the column and in the pixel. The differential readout approach eliminates FPN from the pixel, especially for readout of small signals. The elimination of fixed pattern noise from the columns can be carried out using a double-delta sampling method, reported elsewhere[1], and will not be dealt with in this paper. Sources of temporal noise include white noise and flicker noise from transistors in the source follower, noise in the MOSFET switch that is sampled on the sample-and-hold capacitors, and the so-called reset noise at the pixel associated with resetting the floating diffusion.

Detection of ultra-low light levels requires reduction of input-referred noise as well as increasing responsivity. Input-referred noise for an APS is defined as the equivalent number of electrons at the sense node that produces a voltage at the output equal to the noise voltage resulting from all the noise sources in the signal chain. If $\langle \Delta v_o \rangle$ is the output r.m.s. noise voltage, and g (V/e) is the charge-to-voltage conversion gain from the sense-node to the output, then input-referred noise $\langle N_n \rangle$ in electrons is expressed as:

$$\langle N_n \rangle = \frac{\langle \Delta v_o \rangle}{q_{elec}\, g}. \qquad (1)$$

Thus, we can reduce the input-referred noise by reducing the voltage noise and by increasing the conversion gain at the pixel.

We define the responsivity of the APS as the output voltage change due to the photon collection and define the fill factor as the ratio of the drawn active area to the pixel area. In the rest of the paper, we investigate how to reduce the input-referred noise and increase responsivity by appropriately sizing the pixel transistors and other signal chain parameters, and by optimizing the optical fill-factor in the pixel.

2.1 Noise analysis of the SF

For a source follower driving a sample-and-hold capacitance with input driven by a voltage source, the white noise is independent of transistor sizes, to the first order, and is solely determined by the size of the S/H capacitance. An approximation for the noise power is given by:

$$\langle v_n^2 \rangle \approx \frac{kT}{C}. \qquad (2)$$

Although this simple approximation is an excellent tool to approximately estimate noise in a wide variety of circuits operating in switched-capacitor environments, it is inadequate for accurately predicting low noise behavior in APS. The main reason for this is that the APS pixel source follower is driven not by a voltage, but by a charge input, since FD shown in Figure 1 converts the photocharge into an equivalent potential change at the input of the source follower. Under such a condition, the capacitive feedback network formed by the gate-to-source capacitance (C_g) of M1 and C_{pd}, the effective capacitance at the sense node can significantly alter the noise in the APS pixel, and its effect must be included to modify equation (2).

2.2 Calculation of the noise sources

Taking into account the effect of the capacitive feedback described in the previous section, the transfer function for a small signal voltage (v_o) at the output of the first source follower with respect to the drain current (i_n) can be written as:

$$\frac{v_o}{i_n} = -\frac{C_{fd}+C_{g1}}{g_{mM1}} \frac{1}{C_{eq1}(1+j\frac{w}{w_{eq1}})} \tag{3}$$

where, $C_{eq1} = \frac{C_{fd}+(1-A)C_{g1}}{A}$; $w_{eq1} = \frac{g_{mM1}C_{eq1}}{C_o(C_{g1}+C_{fd})}$; $C_o = C_{SHS} + C_L$, C_{SHS} is the sample-and-hold capacitance, C_L is the line capacitance, C_{fd} is the floating diffusion capacitance and C_{g1} is the capacitive coupling of the pixel SF. The transistor M1 transconductance is g_{mM1}.

Using the transfer function given in (3), and using both white and flicker noise power spectral densities for M1 and M2, an expression for the mean square voltage at the output of the first source follower can be written as:

$$\left\langle v_0^2 \right\rangle = 2kT\left[(1+\beta_1)\frac{C_{fd}+C_{g1}}{C_{eq1}C_o^{'}} + \left(\frac{C_{fd}+C_{g1}}{C_{eq1}}\right)^2 S_{vi}(0)(1+\beta_1^2)\ln(1+\pi\omega_{eq1}\tau/2)\right], \tag{4}$$

where $C_o^{'} = C_o + C_s$, and $C_s = \frac{C_{fd}C_{g1}}{C_{fd}+C_{g1}}$. The first term denotes the white noise contribution and the second term is due to the flicker noise. In the above equation, the following terms are deifined as: $\beta_1 = \frac{g_{mM2}}{g_{mM1}}$, g_m being the transconductance of the specific transistors; $S_{vi}(0) = \frac{K_n}{W_{m1}L_{m1}}$, W_{m1} and L_{m1} are M1 transistor width and length and $K_n = \frac{K_{F_n}}{\mu_n C_{ox}^2}$. K_{Fn} is the device independent flicker noise coefficient, and τ is the separation between the sampling times. The switch drives only the sample and hold capacitance, and the noise added by this MOSFET assumes the familiar kT/C expression:

$$\left\langle v_0^2 \right\rangle = \frac{2kT}{C_{SHS}} \tag{5}$$

In both the above-mentioned expressions, the pre-factor of 2 reflects the effect of the differential readout comprising of two samples.

For the second SF, we ignore the capacitance feedback factor and the flicker noise. The reason is that the APS is customarily designed with a large aspect ratio and C_{SHS}, so that the both feedback factor and flicker noise are negligibly small. Therefore, we obtain:

$$\langle v_o^2 \rangle = kT(1+\beta_2)\frac{1}{C_L'}, \tag{6}$$

where $\beta_2 = \frac{g_{mMp2}}{g_{mMp1}}$, $C_L' = C_{Lout} + C_s'$; C_{Lout} is the output capacitance, and $C_s' = \frac{C_{SHS}C_{g2}}{C_{SHS}+C_{g2}}$ where C_{g2} is the gate-to-source capacitance of the second SF.

Assuming that all the noise are independent of each other, and after adding appropriate voltage gain terms, the total noise at the output of the second source follower is:

$$\langle v_0^2 \rangle = 2kT\left[A_2^2\left\{(1+\beta_1)\frac{C_{fd}+C_{g1}}{C_{eq1}C_o'} + \frac{1}{C_{SHS}}\right\} + (1+\beta_2)\frac{1}{C_L'} + A_2^2\left(\frac{C_{fd}+C_{g1}}{C_{eq1}}\right)^2 \frac{S_{vi}(0)}{kT}(1+\beta_1^2)\ln(1+\pi\omega_{eq}\tau/2)\right] \tag{7}$$

where A_2 is the gain of the second source follower. By dividing the output noise by the conversion gain given by $g = \frac{q_{elec}A_2}{C_{eq1}}$; the input-referred noise can be expressed as:

$$\langle N_n^2 \rangle = \frac{2kT}{q_{elec}^2}\left[(1+\beta_1)\frac{C_{fd}+C_{g1}}{C_o'}C_{eq1} + \frac{C_{eq1}^2}{C_{SHS}} + \frac{C_{eq1}^2}{A_2^2}(1+\beta_2)\frac{1}{C_L'} + (C_{fd}+C_{g1})^2\frac{S_{vi}(0)}{kT}(1+\beta_1^2)\ln(1+\pi\omega_{eq}\tau/2)\right]. \tag{8}$$

In addition to the temporal noise from the signal chain electronics and pixel source follower, there is also the reset noise associated with resetting the floating diffusion. The differential readout approach explained earlier eliminates reset noise in the PG APS by allowing readout of two samples from the same frame. This is not possible with a PD APS; thus, differential sampling readout in a PD APS does not eliminate reset noise, although flicker noise and offsets are suppressed. For PD APS, therefore, there is the additional kTC reset noise term that must be added to the expression in (8). The reset noise term is given by:

$$\langle N_n^2 \rangle = \frac{2kT}{q_{elec}^2}\left[C_{pd} + (1-A_1)C_{g1}\right]. \tag{9}$$

2.3 Theoretical expectations

We note that the white noise expressions presented above exhibit the expected kT/C trend: a reduction in the sample-and-hold capacitance (C_{SHS}) increases the total noise. On the other hand, the flicker noise is independent of the C_{SHS}. This is because C_{SHS} does not determine the bandwidth for flicker noise; the CDS transfer function limits the effective bandwidth for flicker noise to

a much smaller value. Since the white noise component reduces with increasing C_{SHS}, the relative effect of the flicker noise is more dominant for APS designed with smaller C_{SHS}.

The white noise terms in equations (7) and (8) differ from usual kT/C expression because of the capacitive feedback. The capacitive feedback effectively reduces the output resistance of the source follower for the noise, thereby increasing the noise compared to the situation where it is driven by a voltage input. Compared to noise for a voltage mode input, the charge driven source follower can be said to have an "excess" noise, which we quantify in form of a noise multiplier ($\alpha_{M,}$):

$$\alpha_M = \frac{\langle N_n^2 \rangle_{charge-in}}{\langle N_n^2 \rangle_{voltage-in}} = \frac{(C_{fd} + C_g)A}{[C_{fd} + (1-A)C_g]} = \frac{A(1+\frac{C_g}{C_{fd}})}{1+(1-A)\frac{C_g}{C_{fd}}}. \qquad (10)$$

Figure 2 shows the dependence of α_M on the capacitance ratio $\frac{C_g}{C_{fd}}$ and on the amplifier gain A. For the PG case, where the floating diffusion capacitance is minimized, this ratio can be as high as 10, making the effect of the capacitive feedback an important one, especially for low noise design.

Eqns. (7) and (8) indicate that the pixel source follower size affects both the output voltage noise and the input-referred noise in contradictory ways. The white noise tends to increase with increasing source follower area due to the capacitive feedback, while the flicker noise is reduced due to a corresponding reduction in the $S_{vi}(0)$. For large input FET area, the reduction in flicker noise is not as dramatic as it is for smaller geometries, since at larger geometries, the capacitive feedback effect is more prominent. Thus, for a given technology, there is an optimum point in the area domain. This is also evident from Figure 3, which plots the flicker noise, the white noise, and the total noise as a function of area for C_{SH} of 5pF ($S_v(0)$ was taken to be 20 μV/rt-Hz for the smallest size transistor). Our simulations show that this optimum point for the lowest noise lies near the smallest size transistor that can be made in a given technology. Thus, in general, APS noise increases with increasing area of the source follower transistor, contrary to the commonly held notion.

Eqns (7) and (8) indicate that the larger the aspect ratio of the source follower transistor, the smaller the noise. This results from the fact that a larger aspect ratio of the source transistor reduces the effective contribution from the load transistor, thereby reducing overall noise. Figure 4 shows the dependence of the output noise voltage (Vnoise) and the noise equivalent electrons (NEQ) on the SF area and on the aspect ratio. In general, we see a decrease in both the noise voltage and the NEQ with a reduction of the source follower area. However, the decrease is more noticable in the NEQ plot, since a smaller source follower area simultaneously decreases the voltage noise and increases the conversion gain.

3. Experimental results

To verify our conclusions experimentally, we fabricated two test chips, using a 1.2μm CMOS process. The first test chip had 32X32 pixels, with PG designs meant to test the noise dependence on the SH capacitors and the SF parameters in the pixel. The

capacitors changed at the pixel output, where, for every five columns, the capacitance value changed between 0.5pF and 5pF. In addition, the pixels in each row had different SF sizes. The input SF aspect ratio varied between 6/2 and 17/2. The pixel size was kept constant throughout the chip at 20.4 μm. The second test chip was designed to test the optimum pixel parameters for the highest optical response. This sensor had both designs: PD and PG. It consisted of blocks of 4X4 pixels with different active area shapes and sizes and different SF aspect ratios and sizes. The fill factor changed from 6% to 55%, for example, in the PD design. Figure 5 shows pictures of the test chips.

Figure 6 shows the graphs of predicted and measured noise as a function of the input FET area and aspect ratio for two values of C_{SHS}. The kinks in the theoretical curves occur because in the test chip, both area and aspect ratio vary simultaneously. In fact, the kinks result where two different input transistors have the same area, but different aspect ratio. The smaller aspect ratio always leads to higher noise, as to be expected from the theory presented in the previous section. We also note that the larger the sample-and-hold capacitance, the smaller the noise; and the reduction in noise follows an approximate inverse relationship with C_{SHS}, especially for larger aspect ratios. In general, the measured data is in reasonably good agreement with the predicted results. Figure 7, which is the plot of the measured input-referred noise as a function of the input transistor area indicates that the input-referred noise decreases with reduced input transistor area, with the smallest devices resulting in a read noise less than 5 electrons r.m.s. The test results also confirm what the simulation shows: smaller noise requires larger C_{SHS}, smaller transistor area, and larger aspect ratio.

Figure 8 shows the conversion gain (cg) results for PD different designs. There were four main design categories which differed by the shape of the active area (L versus rectangular) and SF parameters. For each design there were several groups of 4x4 blocks of pixels which differed by fill factor. For example, the fill factor increased from 6% in the first group up to 55% in the eighth group in the second pixel design (PD2). The average cg result for PD was 3uV/e. We observed, as expected, a decrease in cg with an increase in the fill factor. We look next at the total responsivity of the pixels to uniform light. Figure 9 plots the second test chip results for the PD and PG pixels. For the PGs, we expect a larger responsivity for a larger fill factor, because a larger signal is collected and the conversion gain should be fairly constant. For the PDs, we expect the responsivity to remain fairly constant when the active area changes, because it simultaneously reduces the conversion gain (as seen in figure 8) through an increase in the PD capacitance, and increases the collected signal. However, from Figure 9, we see an unexpected inverse relationship between fill factor and responsivity. The reason we find for this relation is that the optical response from the "dead" region (i.e., outside the active area) in the pixel increases the collected optical signal without a corresponding increase in capacitance.

4. Conclusions

We have presented a detailed model of the noise performance of APSs. We have analyzed the effects of sizing the input transistor and sample and hold capacitors, and have presented test results that indicate that smallest noise is achieved with larger sample-and-hold capacitors, smaller input transistor area, and larger aspect ratios. We have reported less than 5 electrons read noise from an APS test chip. With another test chip, designed to investigate enhancement of optical response, we have shown

that a smaller pixel fill-factor leads to larger optical responsivity for photodiode type APS. These findings, we believe, will pave the way for future scaled versions of high resolution APS chips for scientific and commercial use.

5. Acknowledgments

The authors thank the JPL Advanced Imager Technology Group for the help during various stages of this work, and Mr. Craig Staller especially for the help with the testing. O.Y-P. appreciates the support of the National Research Council Research Associateship Program and also thanks Mrs. Lyn Dupre' for editing the manuscript.

The research described in this paper was performed by the Center for Space Microelectronics Technology, Jet Propulsion Laboratory, California Institute of Technology, and was sponsored by the National Aeronautics and Space Administration, Office of Space Access and Technology.

6. References

[1]Mendis S. M., Kemeny S. E., Gee R. C., Pain B., Staller C. O., Kim Q. and Fossum E. R., "CMOS active pixel image sensors for highly integrated imaging systems", IEEE J. Solid State Circuits, Vol. 32, No. 2, pp. 187-197, February 1997.

[2]Mendis S. M., Kemeny S. E., and E.R. Fossum, A 128x128 CMOS active pixel image sensor for highly integrated imaging systems, in International Electron Devices Meeting (IEDM) Technical Digest, pp. 583-586, Washington D.C. December 5-8 1993.

[3]Dickinson A., Ackland B., Eid E., Inglis D., and Fossum E. R., "A 256X256 CMOS Active Pixel Image Sensor with Motion Detection", 1995 IEEE International Solid State Circuits Conference, Digest of Technical Papers, pp. 226-227, San Francisco CA, February 1995.

[4]Panicacci R., Pain B., Zhou Z., Nakamura J., Fossum E. R., "Progress in voltage and current mode on-chip analog-to-digital converters," Proc. SPIE 2654, pp. 63-71, 1996.

[5]Panicacci R. , Kemeny S. E., Matthies L. H., Pain B., Fossum E. R., "Multiresolution CMOS active pixel sensor," Proc. SPIE 2654, pp. 72-81, 1996.

[6]Yadid-Pecht O., Pain B., Staller C. O., Clark C., Fossum E. R., "CMOS active pixel sensor star tracker with regional electronic shutter", IEEE J. Solid State Circuits, Vol. 32, No. 2, pp. 285-288, February 1997.

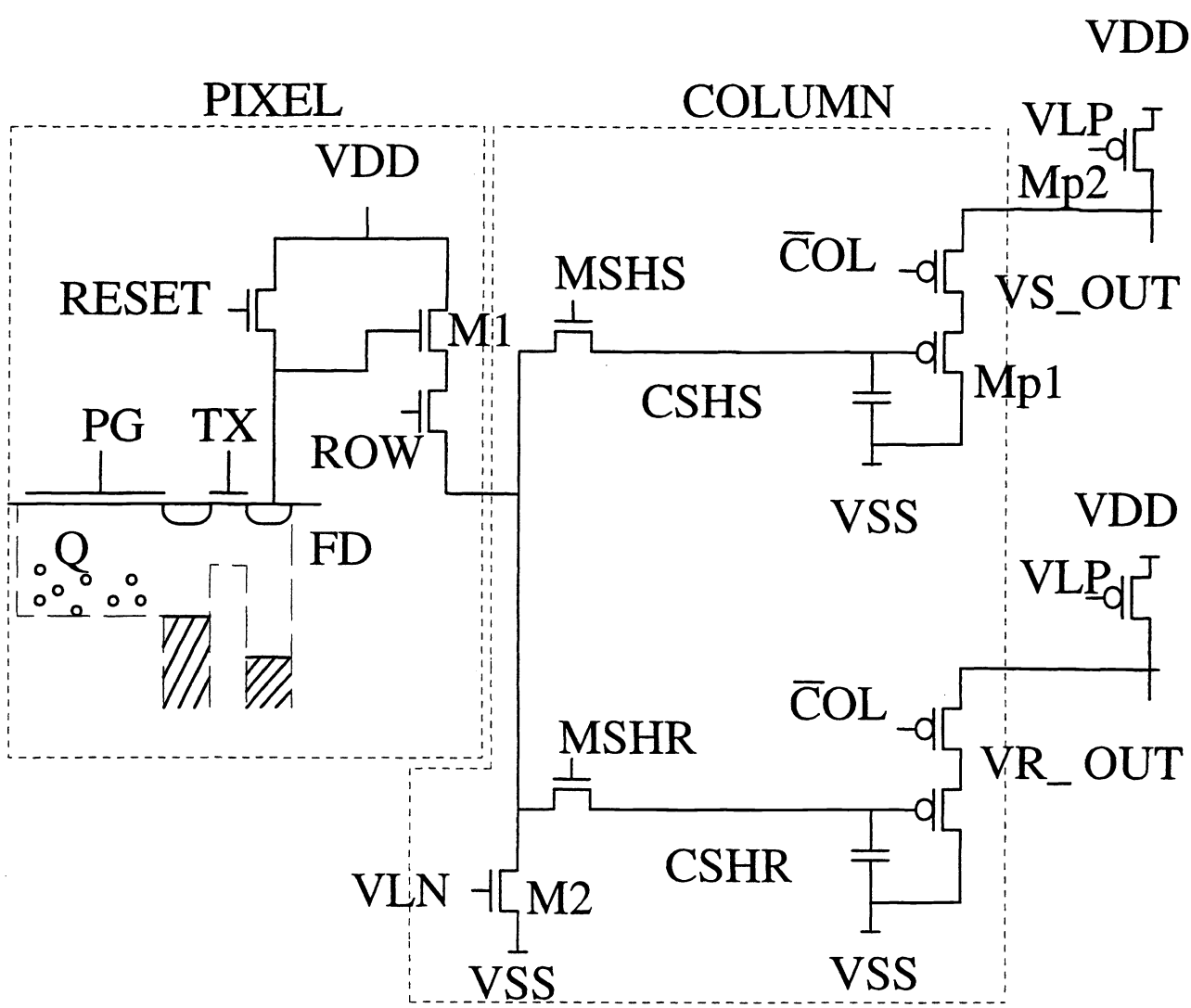


Fig. 1: A photogate CMOS APS pixel and signal chain without the double delta sampling circuit.

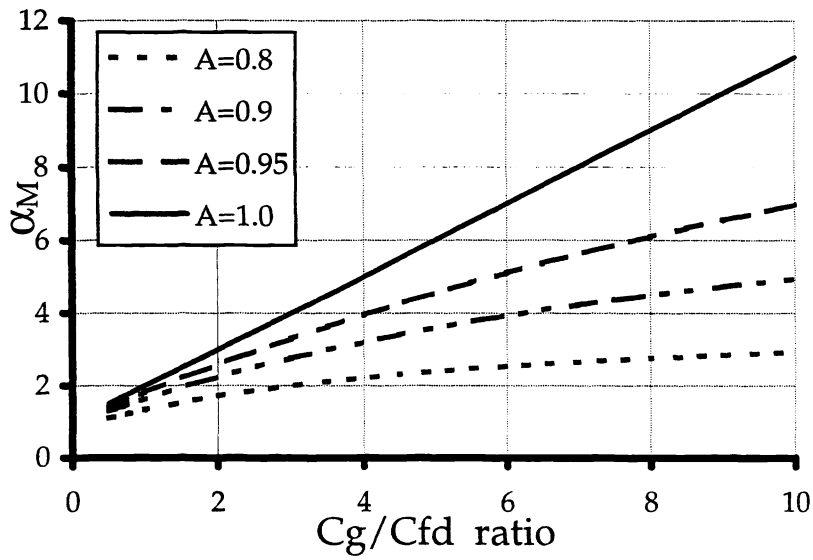


Fig. 2: Excess noise multiplier as a function of area and gain of the SF.

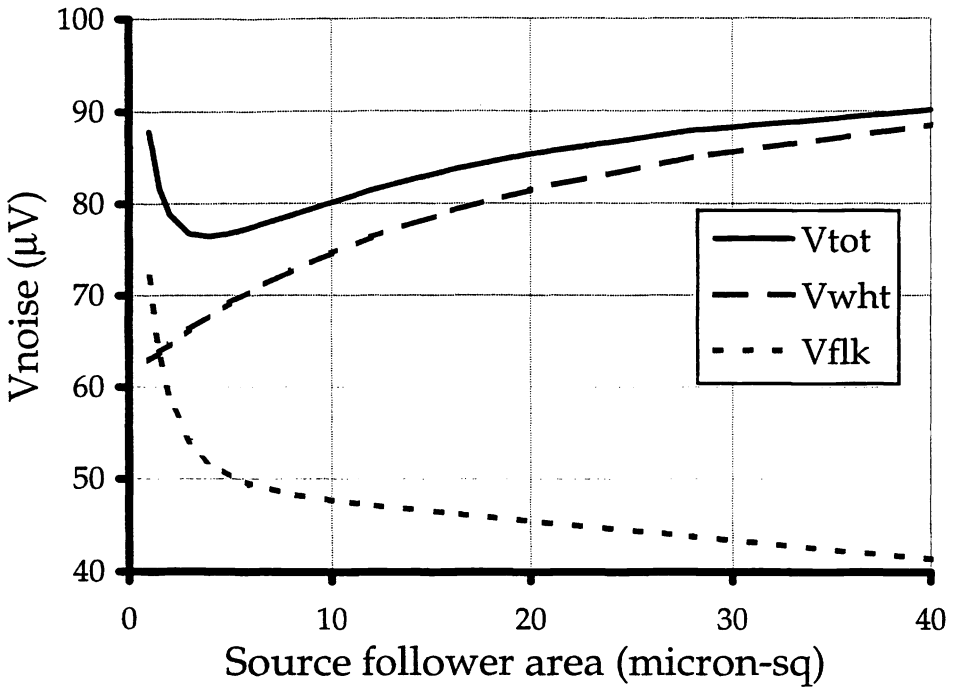
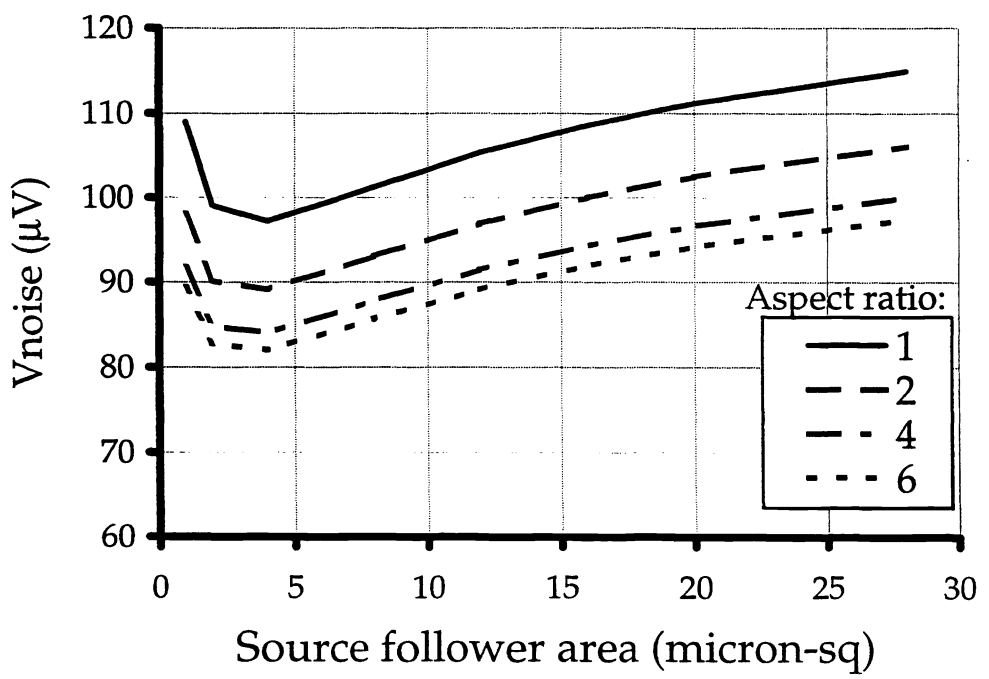
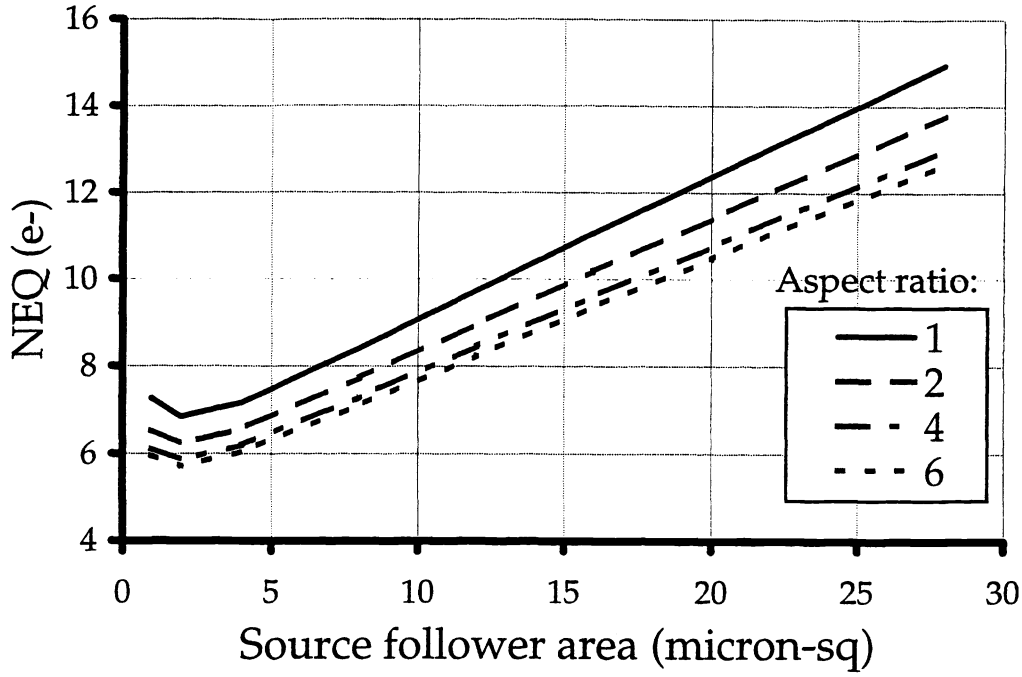


Fig. 3: Flicker noise, white noise and total output noise voltage as a function of the SF area using the parameters specified in the text.



(a)



(b)

Fig. 4: (a) Noise output voltage and (b) input referred electrons as a function of SF area and SF aspect ratio, using the parameters specified in the text.

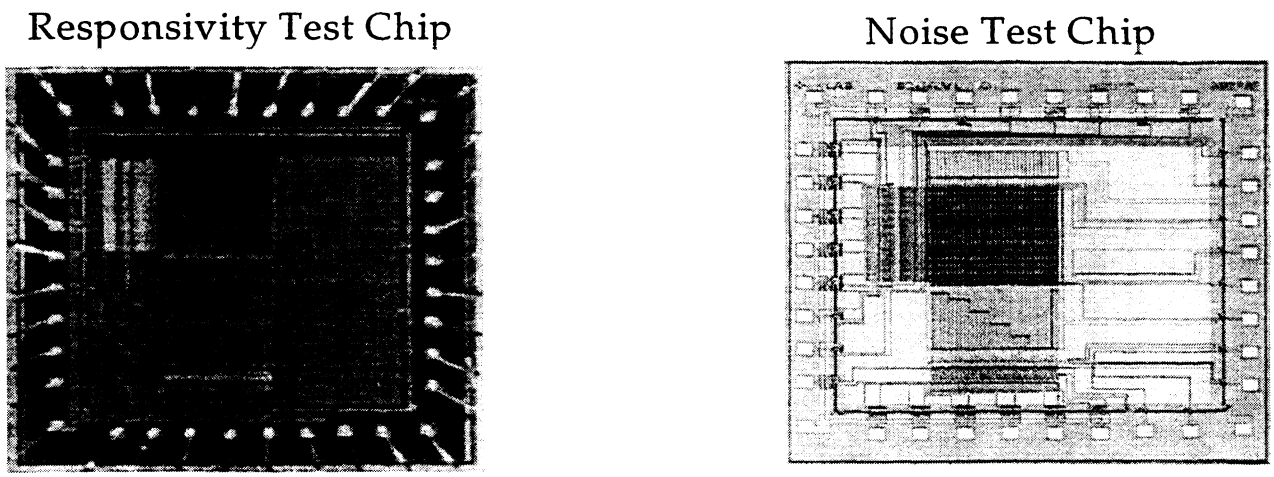


Fig. 5: Test chips for noise and responsivity optimization.

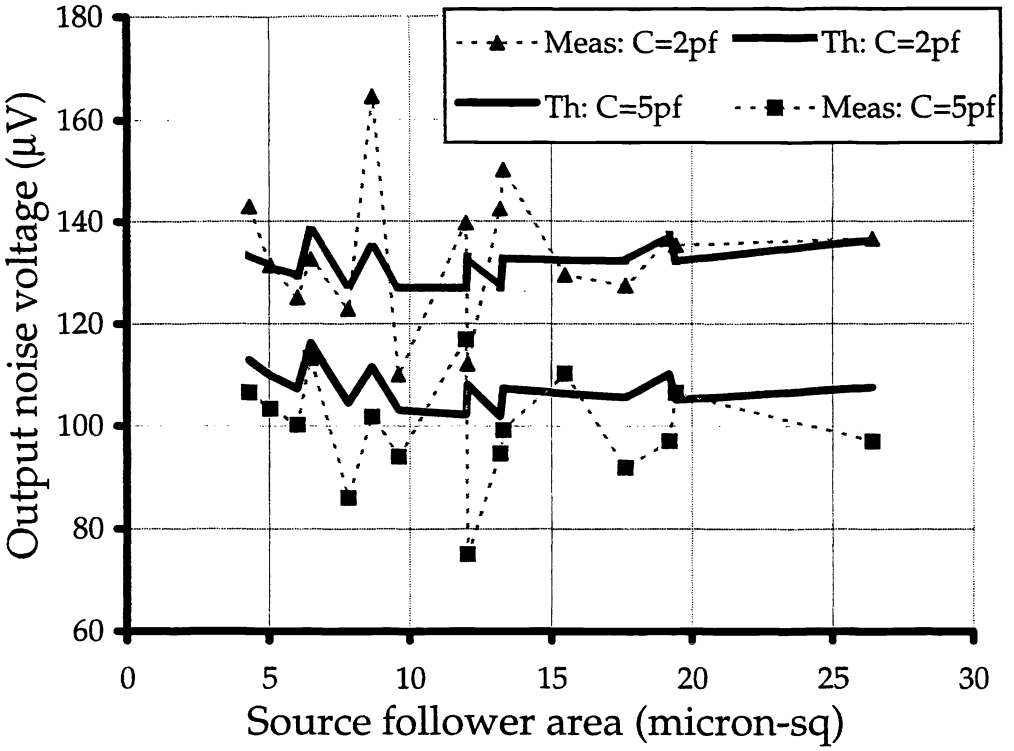


Figure 6: Predicted versus measured data for different C_{SHS}, SF area, and aspect ratios. The parameters are specified in the text, and same design (and conversion gain) was used through this measurement.

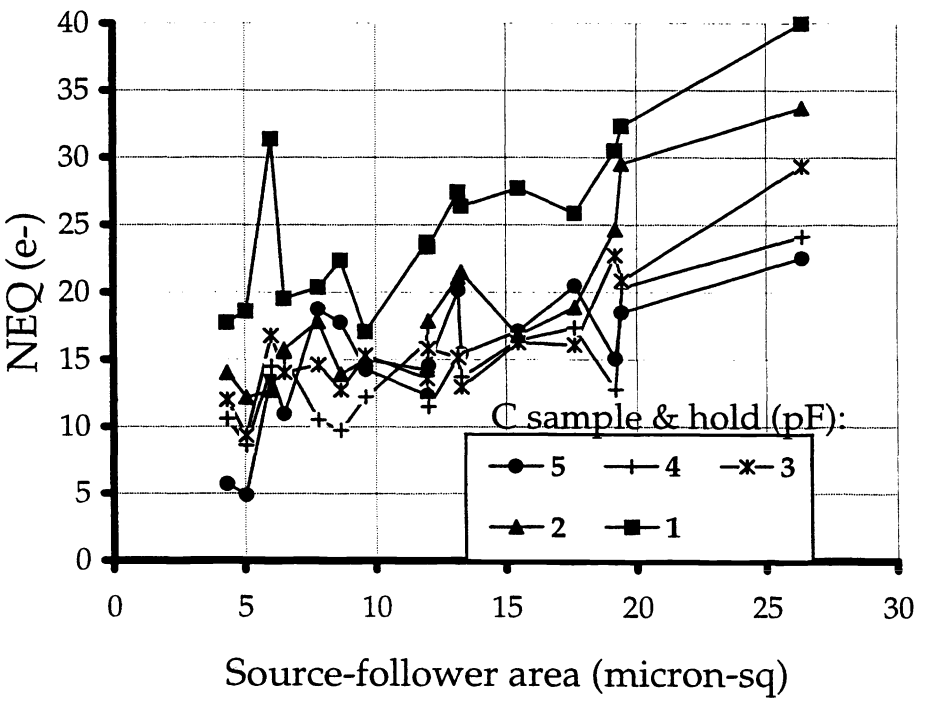


Figure 7: Measured input-referred noise for different input transistor sizes and different C_{SHS}.

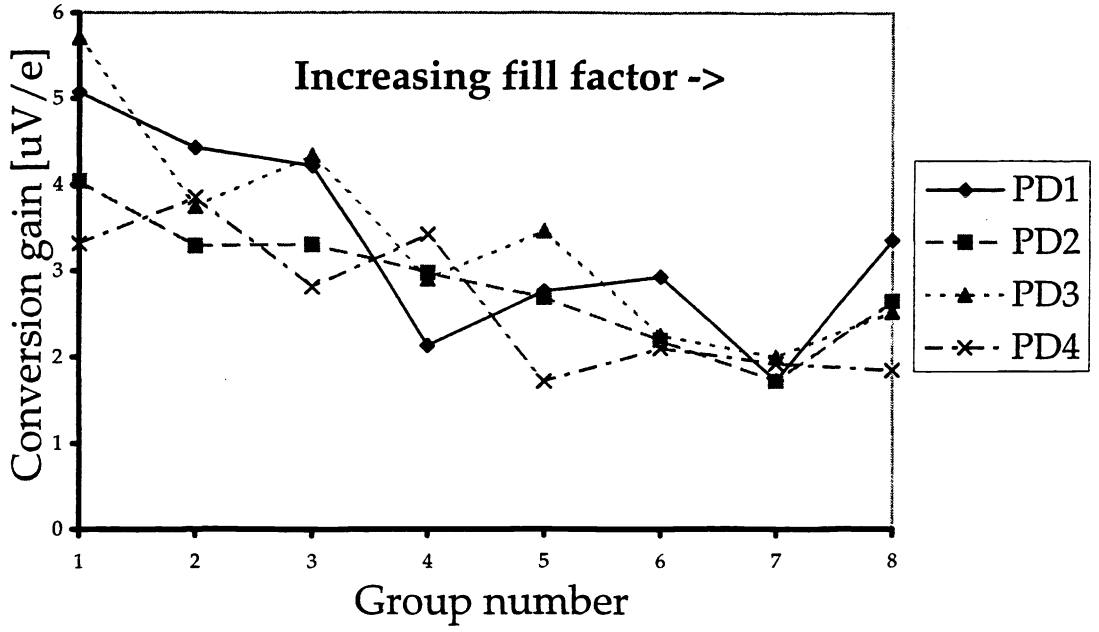


Fig. 8: PD conversion gain results as measured from the second test chip. PD1-PD4 represent different pixel designs, where the fill factor varies for each design. For example, fill factor increases from 6% in the first group up to 55% in the eighth group in the second pixel design (PD2).

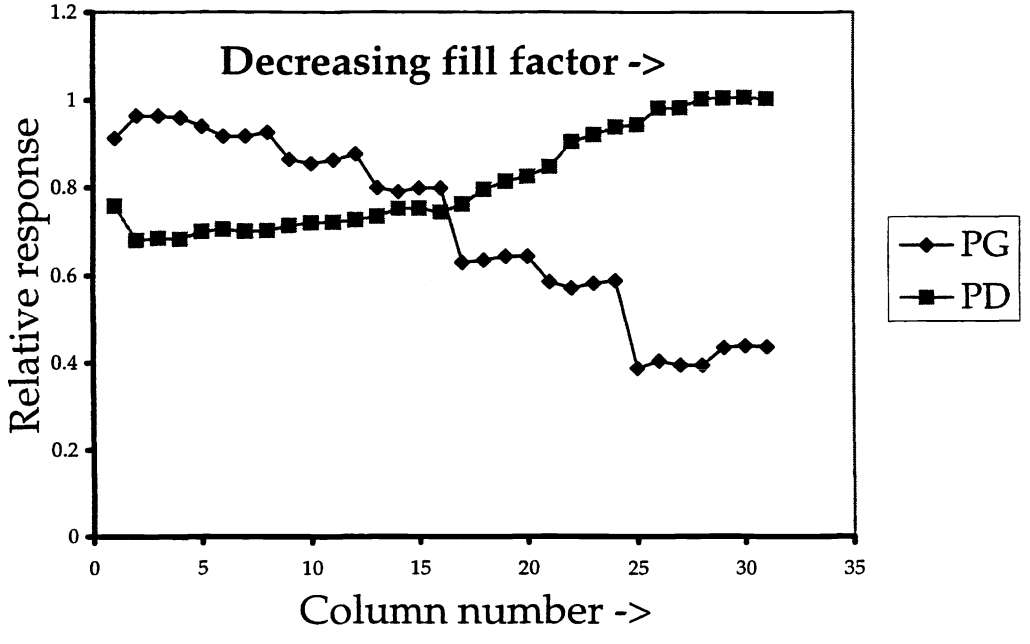


Fig. 9: PD and PG responsivity results as measured from the second test chip for different fill factors of a particular design. For the PD the fill factor decreased from 55% in the first columns to 6% in the last columns.