Politecnico di Milano
MSc Course - MEMS and Microsensors - 2023/2024

E18
Dynamic Range of image sensors

Giacomo Langfelder
giacomo.langfelder@polimi.it

13/12/2023

Problem

A digital still camera features a 3T CMOS active pixel sensor (APS), as the one shown in Fig. 1. Each pixel occupies an area of (3 μm)^2, with a fill factor of 45%. The technology used by the foundry for the production of the sensor allows a maximum voltage of 3.3 V and a minimum size of the transistors (W and L) equal to 150 nm; the oxide capacitance per unit area of a MOS transistor is 2 fF/μm^2. The quantum efficiency of the considered pixel is 0.65. In operation, each photodiode is reverse biased and its depletion region is 1 μm wide. In this configuration, the dark current of the pixel is equal to 0.2 fA.

1. Calculate and graph the dynamic range at the analog output as a function of the integration time, and its value for an integration time of 10 ms.
2. Suitably choose the number of bits for the analog-to-digital converter (ADC) at the output of the sensor.
3. Evaluate the maximum achievable signal-to-noise ratio, plot the SNR vs the photocurrent and cross-verify the dynamic range from that graph.
4. Determine the effects of the variation of the depletion capacitance on the dynamic range.

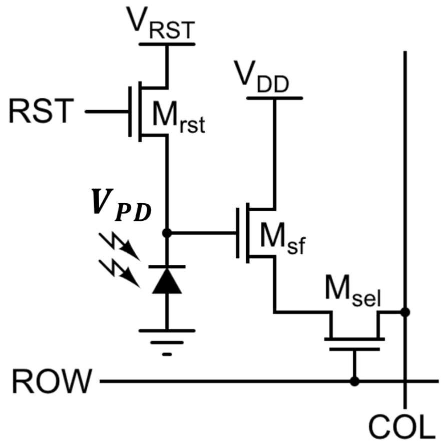

Figure 1: Standard 3T APS.

Part 1: Definition and calculation of the dynamic range

The dynamic range is defined as the ratio between the maximum detectable signal and the minimum detectable signal:

$$DR = \frac{\text{maximum detectable signal}}{\text{minimum detectable signal}}.$$

While the minimum detectable signal is set by noise, the maximum signal is the signal that causes the saturation of the pixel, i.e. the signal for which the output voltage would either rise above or drop below the supply rails of the circuit.

A camera with a good dynamic range is capable to simultaneously distinguish details both in bright and in dark areas of the same scene. Some scenes have a wide dynamic range, meaning that there is a significant difference in brightness between the shadows and the highlights (e.g. shooting a silhouette at sunset), while others have a much narrower range of brightness levels. When you take a picture, there are actually two dynamic ranges to consider: the dynamic range of the scene you're photographing and the dynamic range of the imaging sensor. All pixels inside your digital camera can only record a fixed dynamic range (in a single exposure). As long as the difference in brightness between the darkest and lightest areas of a scene fall within this dynamic range, you'll be able to record detail in both areas simultaneously. For example, if a camera sensor has a certain dynamic range and the difference between the shadows and highlights of the scene is lower, then you will be able to capture detail in all areas of the scene. However, if the dynamic range of the scene exceeds that of the sensor, you'll end up with a picture where the shadows are completely black or where the highlights have 'blown' and become totally white - and sometimes both.

Regarding the integration capacitance, with the given data, the photodiode capacitance turns out to be

$$C_{PD} = \frac{\epsilon_{Si} A_{pixel} FF}{x_{dep}} = 0.42\,\text{fF}.$$

We know that the integration capacitance is given by the parallel (the sum) of the photodiode intrinsic capacitance and the parasitic capacitance from the integration node to ground:

$$C_{int} = C_{PD} + C_P.$$

The main contribution to this parasitic capacitance is given by the input capacitance of the source follower. Let's assume a minimum-size transistor: $W \simeq L \simeq 150$ nm; in order to have high fill factors, transistors within the pixel will be likely designed with minimum lengths. In this situation (rough approximation),

$$C_{SF} \simeq C'_{ox} WL \simeq 2\,\text{fF}/\mu\text{m}^2 \cdot 0.15\,\mu\text{m} \cdot 0.15\,\mu\text{m} = 0.045\,\text{fF},$$

which is lower than the photodiode capacitance. Hence, the assumption $C_{int} \simeq C_{PD}$ is valid (note, however, that this is not always the case).

We can make a couple of reasonable approximations, to estimate the dynamic range:

- when estimating the maximum signal, we suppose that the photocurrent will be much larger than the dark current: $i_{ph} \gg i_d$;
- when estimating the minimum detectable signal, we suppose that the photocurrent will be much smaller than the dark current: $i_{ph} \ll i_d$, i.e. dark shot noise dominates over signal shot noise.

The *minimum detectable signal*, $N_{ph,min}$, is the one for which the signal to noise ratio equals one, i.e. when $SNR = 1$. We do not consider here the ADC quantization noise, which will be later designed in such a way that its noise contribution can be neglected with respect to the dark shot noise and reset noise. We get:

$$\sigma_{dark,N} = \frac{\sqrt{q i_d t_{int}}}{q} = 3.5\,\text{electrons}.$$

and

$$\sigma_{reset,N} = \frac{\sqrt{kTC_{PD}}}{q} = 8.3\,\text{electrons}.$$

Hence:

$$N_{ph,min} = \sigma_{tot,N} \simeq \sqrt{\sigma^2_{dark,N} + \sigma^2_{reset,N}} = 8.9\,\text{electrons}.$$

We should now determine the maximum signal, in terms of charge, that can be accumulated by the pixel. The total charge is the photocurrent integrated on the capacitance of the photodiode. The capacitance, in turn, depends on the dopings (set by the process used in sensor manufacturing), the pixel area (set during pixel layout), and the bias voltage (an operating parameter). The size of this capacitance creates a limit on the total amount of charge a pixel can hold during integration. When considering the case of the 3T APS, assuming that all gate-source DC voltages can be neglected, as the maximum variation of the photodiode voltage is V_{DD}, saturation is avoided as long as

$$\frac{Q_{ph}}{C_{PD}} < V_{DD}.$$

The *full-well charge*, FWC, or the *maximum detectable signal*, $Q_{ph,max}$, can be thus estimated as the product of C_{PD} and the maximum voltage swing across the photodiode, V_{DD}:

$$Q_{ph,max} = V_{DD} C_{PD} = 3.3\,\mathrm{V} \cdot 0.42\,\mathrm{fF} = 1.38 \times 10^{-15}\,\mathrm{C},$$

i.e.

$$N_{ph,max} = \frac{Q_{ph,max}}{q} = 8650\,\mathrm{electrons}.$$

Said in other words, the pixel is initially reset at the supply voltage $V_{RST} = V_{DD}$; during the integration phase, the voltage linearly decreases down to a certain value:

$$V_{PD}\left(t_{int}\right) = V_{DD} - \Delta V_{PD} = V_{DD} - \frac{(i_{ph} + i_d)}{C_{int}} t_{int}.$$

In this condition, if the diode voltage falls below 0 V, the pixel saturates. When the pixel saturates - see e.g. the blue (the steepest) curve in Fig. 2 - the output voltage is no longer a valid signal.

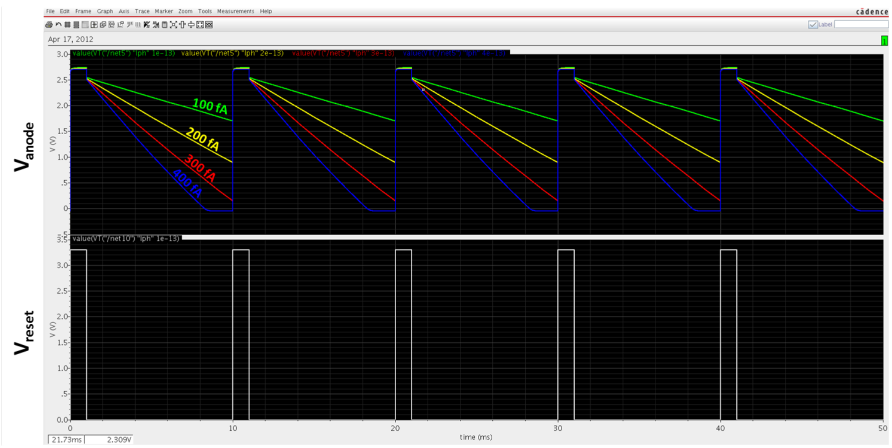

Figure 2: Anode voltage of the photodiode in a 3T topology. The four curves correspond to increasing photocurrents. The blue one experiences pixel saturation.

The dynamic range can now be easily calculated:

$$DR = \frac{Q_{ph,max}}{Q_{ph,min}} = \frac{1.38 \times 10^{-15}\,\mathrm{C}}{1.43 \times 10^{-18}\,\mathrm{C}} = \frac{N_{ph,max}}{N_{ph,min}} = \frac{8650\,\mathrm{electrons}}{8.9\,\mathrm{electrons}} = 965 = 59.7\,\mathrm{dB}.$$

We can write the complete expression of the dynamic range of our sensor:

$$DR = \frac{Q_{ph,max}}{Q_{ph,min}} = \frac{V_{DD}C_{PD}}{\sqrt{qi_d t_{int} + kTC_{PD}}} = \frac{V_{DD}\dfrac{\epsilon_{Si}A_{pixel}FF}{x_{dep}}}{\sqrt{qJ_d A_{pixel}FFt_{int} + kT\dfrac{\epsilon_{Si}A_{pixel}FF}{x_{dep}}}}.$$

Some considerations (see Fig. 3):

- the dynamic range is independent of the signal, i.e. it is an intrinsic property of the sensor; note that the dynamic range is the same for all the pixels, while the signal-to-noise ratio can change across different pixels;
- the dynamic range depends on the integration time: the longer the exposure time, the poorer the dynamic range; the maximum dynamic range is thus obtained for low integration times, when $qJ_d A_{pixel}FFt_{int} \ll \sigma^2_{reset}$; further decreasing the integration time has no effect on the dynamic range;
- the dynamic range improves with the square root of the area, i.e. bigger pixels with larger overall area and fill factor are better;
- the dynamic range is technology dependent, as the dark current density J_d depends on the quality of the silicon substrate;
- the dynamic range improves when increasing the supply voltage.

Fig. 3 reports the graphical solution of the dynamic range as a function of the integration time, for three different diode areas.
The just calculated dynamic range (59.7 dB) is not the maximum dynamic range of the sensor, which can be calculated for short integration times, i.e. when assuming the dark current shot noise negligible, i.e. when reset noise dominates:

$$DR_{max} = DR\left(t_{int} = 0\right) = \frac{V_{DD}\dfrac{\epsilon_{Si}A_{pixel}FF}{x_{dep}}}{\sqrt{kT\dfrac{\epsilon_{Si}A_{pixel}FF}{x_{dep}}}} = \frac{V_{DD}}{\sqrt{\dfrac{kT}{C_{PD}}}} = \frac{8650\,\text{electrons}}{8.3\,\text{electrons}} = 1050 = 60.4\text{dB}.$$

For comparison, the dynamic range of the human eye is around 90 dB, while real world scenes have dynamic ranges far greater than 100 dB. This is the reason why a lot of research is put into finding strategies for high dynamic range (HDR) digital imaging.

Part 2: Choice of the ADC

The analog-to-digital converter (ADC) is used to convert the acquired analog data into the digital domain.

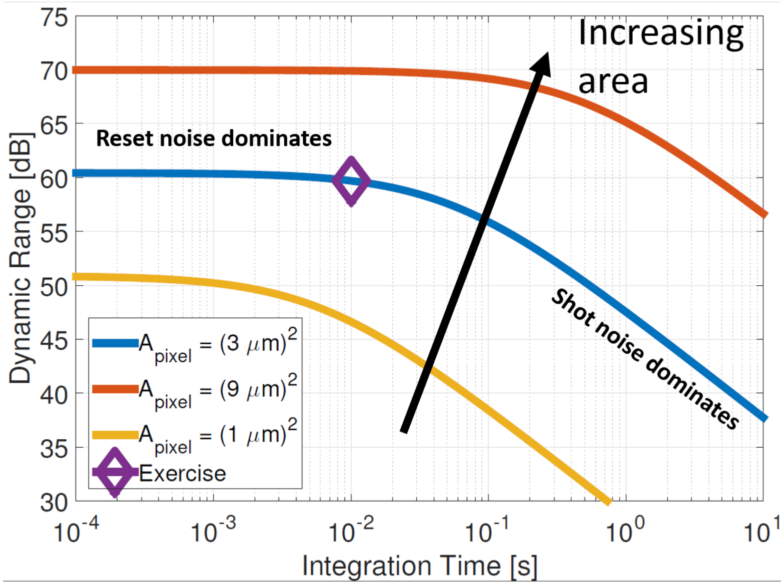

Figure 3: Dynamic range as a function of the integration time, for different pixel areas.

The maximum digital level of an analog-to-digital converter should correspond to the maximum analog voltage swing, i.e.:

$$V_{ref,ADC} = 2^{N_{bit}} \cdot LSB = V_{DD},$$

where N_{bit} is the number of bits of the ADC and LSB is the least significant bit, expressed in volts:

$$LSB = \frac{V_{ref,ADC}}{2^{N_{bit}}},$$

An ADC with a high number of bits enables low quantization noise, but the raw output requires more space in the storing memory. A possible way to choose the ADC resolution is to make the quantization noise just lower than the analog noise given by the pixel itself. As minimum noise occurs at short integration times, where reset noise dominates, we set

$$\sigma_{quant,V} = \frac{LSB}{\sqrt{12}} \leq \sigma_{reset,V}.$$

Hence:

$$\frac{\frac{V_{ref,ADC}}{2^{N_{bit}}}}{\sqrt{12}} = \frac{\frac{V_{DD}}{2^{N_{bit}}}}{\sqrt{12}} \leq \sigma_{reset,V}, \qquad N_{bit} > \log_2\left(\frac{V_{DD}}{\sqrt{12}\sigma_{reset,V}}\right) = 8.24.$$

We can then choose an ADC with 9 bits. We can calculate the corresponding quantization noise in terms of number of electrons, as

$$\sigma_{quant,N} = \frac{\frac{V_{ref,ADC}}{2^{N_{bit}}}}{\sqrt{12}} C_{PD} \frac{1}{q} = 4.9 \text{ electrons}.$$

As desired, this noise contribution is lower than the previously-calculated reset noise.

Part 3: Maximum signal-to-noise ratio

We are now interested in finding the maximum signal-to-noise ratio that can be obtained, assuming a certain photocurrent for which the pixel gets close to saturation. We can write the expression of the signal-to-noise ratio for the maximum signal $Q_{ph,max}$:

$$SNR_{max} = \frac{Q_{ph,max}}{Q_{ph,min}} = \frac{Q_{ph,max}}{\sqrt{q\,(i_{ph,max} + i_d)\,t_{int} + kTC_{PD}}}.$$

Since we are dealing with the maximum SNR, i.e. with high signals, we can reasonably assume that, in this situation, the dominant noise contribution is given by the signal shot noise. Hence,

$$SNR_{max} \simeq \frac{Q_{ph,max}}{\sqrt{q i_{ph,max} t_{int}}} = \frac{Q_{ph,max}}{\sqrt{q Q_{ph,max}}} = \sqrt{\frac{Q_{ph,max}}{q}} = \sqrt{N_{ph,max}}$$

$$SNR_{max} = \sqrt{N_{ph,max}} = \sqrt{8638} = 92.9 = 39.4 \text{ dB}.$$

The maximum signal-to-noise ratio does not depend on the integration time, and is proportional to the square root of the maximum number of electrons $N_{el,max}$ that can be collected by the photodiode. This is an expected result. As signal shot noise dominates, noise distribution follows Poisson statistics: given a certain signal N, its variance is equal to the signal itself, $\sigma_N^2 = N$, i.e. $\sigma_N = \sqrt{N}$.
The maximum signal-to-noise ratio thus increases with the supply voltage V_{DD} and with the diode area (and with the fill factor, as well).
The maximum signal-to-noise ratio does not depend on the integration time.
The complete graph of the signal-to-noise ratio as a function of the photocurrent is represented in Fig. 5. The dynamic range can be evaluated from this graph by considering the *x-axis distance* between the two points were the SNR goes to 0 dB (the first one refers to the minimum detectable signal, the second one refers to pixel saturation, i.e. the maximum signal).

Part 4: Depletion capacitance effect

Up to now, we assumed that the width of the depletion region and therefore the depletion capacitance of the photodiode, on which the charge is integrated, is constant ($x_{dep} = 1$ μm). If the voltage variation $\Delta V_{PD} = (i_{ph} + i_d)\,t_{int}/C_{PD}$ is small, with respect to V_{DD}, this assumption is correct. Otherwise, a more correct model should take into account the dependence of the depletion region width on the reverse bias voltage:

$$x_{dep} = \sqrt{\frac{2\epsilon_0\epsilon_r\,(V_{PD} + V_{bi})}{qN_A}},$$

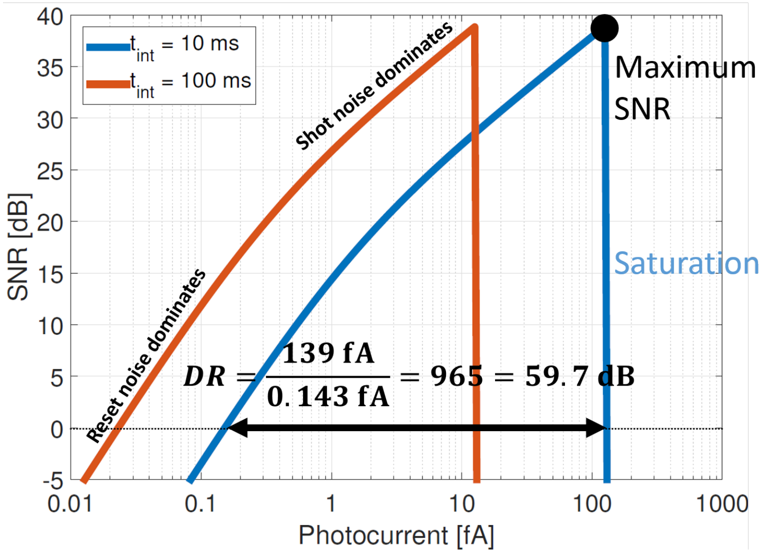

Figure 4: SNR as a function of the photocurrent, for two integration times.

where $V_{PD} = V_{DD} - \Delta V_{PD}$ and V_{bi} is the Silicon built-in voltage (typically about 0.7 V). This results in a photodiode capacitance C_{PD}:

$$C_{PD} = \frac{\epsilon_{Si} A}{x_{dep}} = A \cdot \sqrt{\frac{q \epsilon_{Si} N_A}{2(V_{PD} + V_{bi})}} = \frac{C_0}{\sqrt{1 + \frac{V_{PD}}{V_{bi}}}},$$

where $C_0 = 1fF$ is the photodiode capacitance with $V_{PD} = 0$ V (different from the capacitance computed during reset, with $V_{PD} = V_{DD}$, that we used so far).
We can integrate the capacitor equation to find the exact charge integrated across C_{PD}:

$$Q_{ph,max} = \int_{V_{DD}}^{0} C(V_{PD})\, dV_{PD} = \int_{V_{DD}}^{0} \frac{C_0}{\sqrt{1 + \frac{V_{PD}}{V_{bi}}}}\, dV_{PD}$$

We substitute $x = V_{PD}/V_{bi}$ to compute the integral and we get the final result:

$$Q_{ph,max} = \int_{\frac{V_{DD}}{V_{bi}}}^{0} \frac{C_0 \cdot V_{bi}}{\sqrt{1+x}}\, dx = C_0 V_{bi} \int_{\frac{V_{DD}}{V_{bi}}}^{0} \frac{dx}{\sqrt{1+x}} = C_0 V_{bi} \left[2\sqrt{1+x}\right]_{\frac{V_{DD}}{V_{bi}}}^{0} =$$
$$= 2 C_0 V_{bi} \left(1 - \sqrt{1 + \frac{V_{DD}}{V_{bi}}}\right).$$

$$N_{el,max} = \frac{Q_{ph,max}}{q} = \frac{2 C_0 V_{bi}}{q} \left(1 - \sqrt{1 + \frac{V_{DD}}{V_{bi}}}\right) = -12167\ electrons$$

The result is a negative number, which reflects the fact that the voltage across the diode decreases. We can consider its absolute value.

With this result, we can quantify the percentage error we commit on the maximum charge when considering a constant capacitance equal to $C_{PD,const} = \frac{C_0}{\sqrt{1+V_{DD}/V_{bi}}}$:

$$\epsilon_{lin} = \frac{12167 - 8650}{12167} = 0.289 \approx 30\%,$$

computed with $V_{DD} = 3.3$ V and $V_{bi} = 0.7$ V. This result means that the maximum signal that we computed is actually 70% of the maximum signal a real (non-linear) photodiode would reach. Therefore, without approximation we would get a DR larger by about $20 \cdot \log_{10}(1/0.7) = 3$ dB than what we computed a few pages ago. Bear in mind that the MOS capacitance and parasitic capacitance due to the interconnections, in parallel with the diode, are linear: this mitigates the discrepancy that we have described.

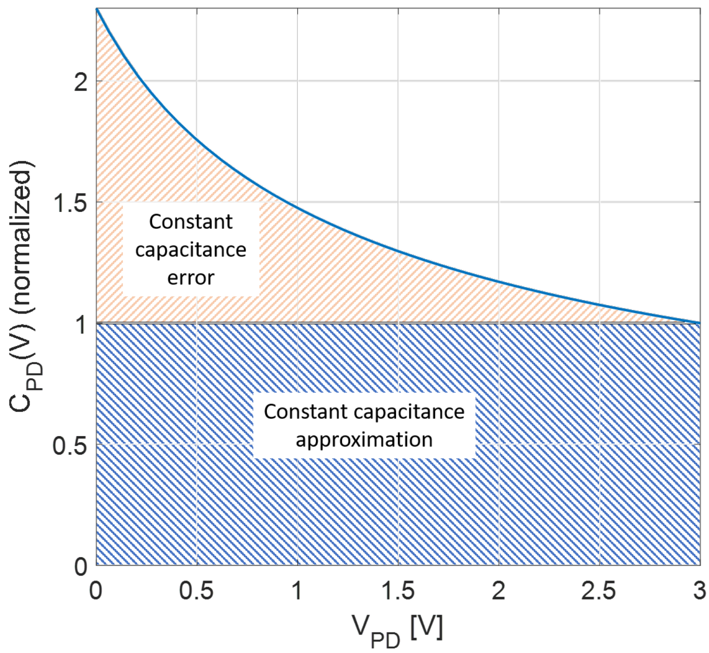

Figure 5: Non-linear $C(V)$ function of a photodiode. The area below the curve is the charge integrated onto the diode capacitance. The blue and orange areas are the charge computed with the constant-capacitance approximation and its error.