Lecture Notes 3
Introduction to Image Sensors

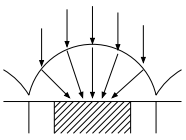
EE 392B
Prof. A. El Gamal

Handout #5
Spring 01

- **CCDs**
  - basic operation
  - well capacity
  - charge transfer efficiency and readout speed
- **CMOS Passive Pixel Sensor (PPS)**
  - basic operation
  - charge to output voltage transfer function
  - readout speed
- **CMOS Active Pixel Sensor (APS)**
  - basic operation
  - charge to output voltage transfer function
  - readout speed
- **Photogate APS**

Preliminaries

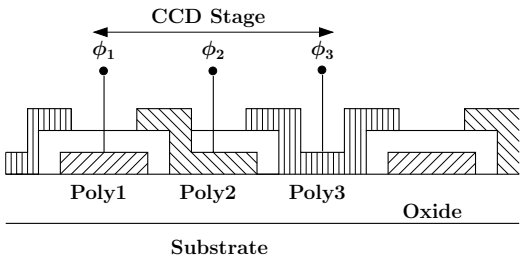
- area image sensor array consists of $n \times m$ pixels, ranging from 320×240 (QVGA) to 7000×9000 (very high end Astronomy)
- each pixel contains a photodetector and devices for readout (capacitors for CCD, MOS transistors for CMOS sensors)
  - pixel size ranges from $15 \times 15 \mu\text{m}^2$ down to $\approx 3 \times 3 \mu\text{m}^2$ (limited by dynamic range and cost of optics)
  - *fill factor* is the fraction of pixel area occupied by the photodetector and ranges from 0.2 to 0.9 — high fill factor is desirable (high well capacity and sensitivity (why?))
  - fill factor can be increased using microlens



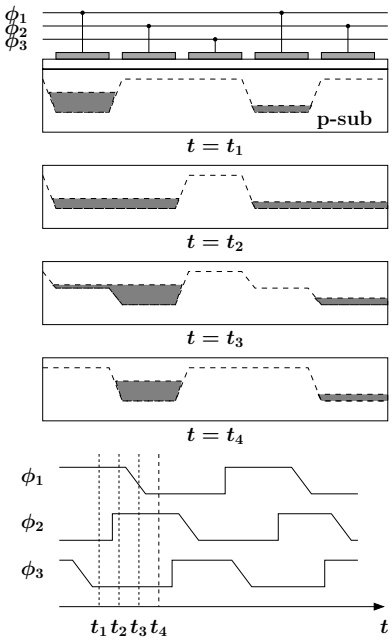
- the primary difference between CCD and CMOS image sensors is the readout architecture
  - for CCDs, charge is shifted out
  - for CMOS image sensors, charge or voltage is read out using row and column decoders — similar to a digital memory (but analog data is read out)
- the readout circuits (including in pixel devices) determine the sensor *conversion gain*, which is the output voltage per electron collected by the photodetector, in μV/electron
- given the sensor spectral response, conversion gain, and area its *sensitivity* measured in V/Lux.s can be determined
- readout speed determines the video *frame rate* that an image sensor can operate at — 30 to 60 frames/s are typical, but lower frame rates are sometimes dictated by the available bandwidth (*e.g.*, PC camera), and higher frame rates are required for many industrial and military applications

CCD Image Sensors

- a CCD is a *dynamic analog (charge) shift register* implemented using closely spaced MOS capacitors clocked using 2, 3, or 4 phase clocks – capacitors operate in deep depletion regime when clock is high
- charge transfer (from one capacitor to the next) must occur at high enough rate to avoid corruption by leakage, but slow enough to ensure high charge transfer efficiency
- a three phase CCD is typically implemented using three polysilicon layers (for clock routing and improving fringing field)

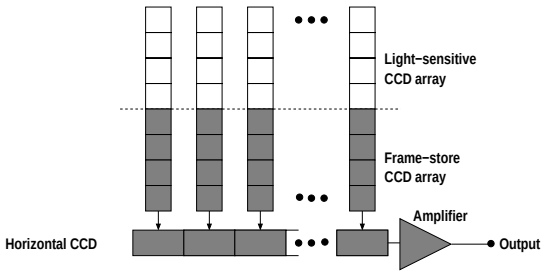


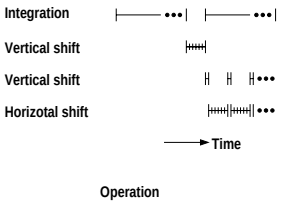
3-Phase CCD Operation



Potential wells and timing diagram during charge transfer

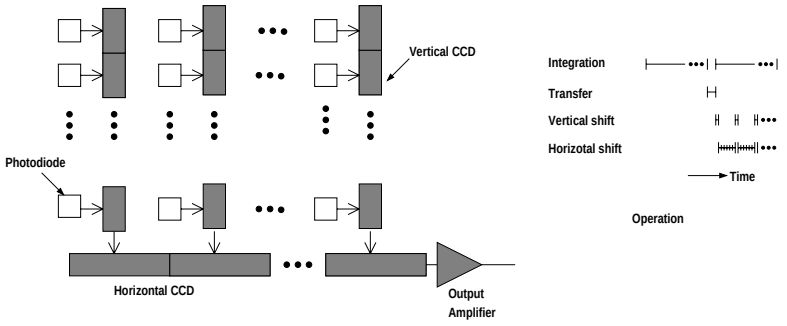
Frame Transfer CCD Image Sensor





- top CCD array used for photodetection (photogate) and vertical shifting
- bottom CCD array optically shielded – used as frame store
- operation is pipelined: data is shifted out via the bottom CCDs and the horizontal CCD during integration time of next frame
- transfer from top to bottom CCD arrays must be done very quickly to minimize corruption by light, or in the dark (using a mechanical shutter)
- output amplifier converts charge into voltage, determines sensor conversion gain

Interline Transfer CCD Image Sensor



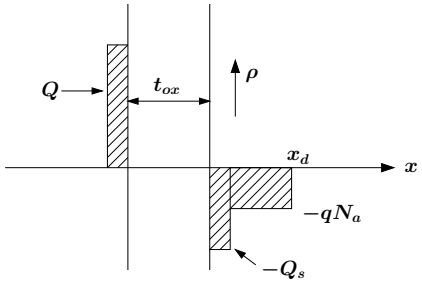
- photodiodes are used
- all CCDs are optically shielded, used only for readout
- collected charge is *simultaneously* transferred to the vertical CCDs at the end of integration time (a new integration period can begin right after the transfer) and then shifted out
- charge transfer to vertical CCDs simultaneously *resets* the photodiodes, (shuttering done electronically for 'snap shot' operation)

Frame Transfer Versus Interline Transfer

- frame transfer uses simpler technology (no photodiodes), and achieves higher fill factor than interline transfer
- interline transfer uses optimized photodiodes with better spectral response than the photogates used in frame transfer
- in interline transfer the image is captured at the same time ('snap shot' operation) and the charge transfer is not subject to corruption by photodetection (can be avoided in frame transfer using a mechanical shutter)
- frame transfer chip area (for the same number of pixels) can be larger than interline transfer
- most of today's CCD image sensors use interline transfer

Well Capacity for CCD

- the well capacity for a CCD is the well capacity for an MOS capacitor operating in the deep depletion regime
- to find the well capacity, consider the charge configuration including signal charge Q_s Col/cm^2



- we modify the equations provided in Appendix III of Lecture notes 2 to get

$$v_G + v_{FB} = \psi_s + \frac{1}{C_{ox}}(qN_a x_d + Q_s),$$

using the relation

$$x_d = \sqrt{\frac{2\epsilon_s}{qN_a} \cdot \psi_s}$$

and the fact that to stay in deep depletion we must have

$$\psi_s > 2\phi_p$$

we get that

$$Q_s < C_{ox} \cdot (v_G - v_T) \text{ Col/cm}^2,$$

where

$$v_T = 2\phi_p - v_{FB} + \frac{1}{C_{ox}}\sqrt{4qN_a\epsilon_s\phi_p}$$

Charge Transfer Efficiency

- the CCD charge *transfer efficiency*, $\eta \leq 1$, is the fraction of signal charge transferred from one CCD stage to the next
- must be made very high (≈ 1) since in a CCD image sensor charge is tranferred up to $n+m$ CCD stages ($3 \times (n+m)$ times for 3-phase CCD)
- Example: consider a 1024×1024 CCD image sensor
  the table lists the charge transfer efficiency η and the corresponding worst case fraction of charge transferred to the output

η	fraction at output
0.999	0.1289
0.9999	0.8148
0.99999	0.9797

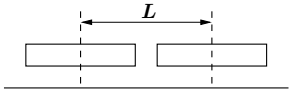
- charge transfer mechanism:
  - most of the charge is transferred very quickly by repulsive force among electrons, which creates self-induced lateral drift
  - remaining charge transferred slowly by thermal diffusion and fringing field
  - transfer efficiency can be acurately estimated using 2 or 3D device simulation
- simple analysis:
  - 99% of charge transferred immediately and remaining 1% transferred slowly by thermal diffusion
  - last 1% accounts for most of the transfer time, and we can write

$$\eta = (1 - 0.01e^{-\frac{t_{stage}}{p\tau}})^p,$$

where t_{stage} is the CCD stage transfer time, p is the number of CCD phases used, and

$$\tau = \frac{4L^2}{\pi^2 D_n},$$

where, L is the center to center distance of adjacent capacitors and D_n is the diffusion constant at the surface



  - given a desired η, we can use the above equation to find a lower bound on transfer time

- more accurate analysis must consider fringing field, which increases η, and the surface trap states, which reduce it
- fringing field is increased by:
  - increasing the gate voltage difference during transfer
  - decreasing L (which also decreases τ)
  - reducing the spacing between capacitors (making it comparable to the oxide thickness) and overlapping the poly gates
  - using low substrate doping

CCD Readout Speed

- CCD imager readout speed is limited mainly by the array size and the charge transfer efficiency requirement
- Example: consider a 1024×1024 3-phase interline transfer CCD image sensor with $\eta = 0.99997$, $L = 4\mu$m, and $D_n = 35\text{cm}^2/s$, find the maximum video frame rate

  transfer time for the horizontal CCD limits the readout speed
  to find the minimum required transfer time per CCD stage, t_{min}, we use the equation

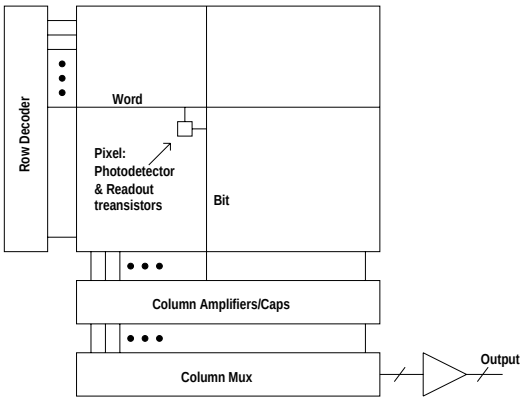
$$\eta = 0.99997 = (1 - 0.01e^{-\frac{t_{min}}{3\tau}})^3,$$

  which gives $t_{min} = 37.8$ns, thus the time required to shift one row out is $37.8\text{ns} \times 1024 = 38.7\mu$s
  ignoring the row transfer time (from vertical CCDs), we get
  minimum frame transfer time of 39.6ms, or maximum video frame rate of 25frame/s
  Note: CMOS image sensors can be much faster than CCDs

Advantages and Disadvantages of CCDs

- Advantages: high quality
  - optimized photodetectors — high QE, low dark current
  - very low noise — no noise introduced during shifting
  - very low fixed pattern noise (nonuniformity) — no FPN introduced by shifting
- Disadvantages:
  - cannot integrate other analog or digital circuits, *e.g.*, for clock generation, control, or A/D conversion
  - highly nonprogrammable, *e.g.*, difficult to implement *window of interest*
  - high power — entire array switching all the time (high C, high V, and high f result in high CV^2f)
  - limited frame rate (for large sensors) due to required increase in transfer speed (while maintaining acceptable transfer efficiency)

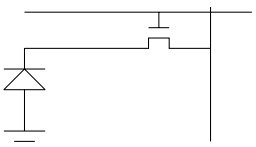
CMOS Passive and Active Pixel Sensors



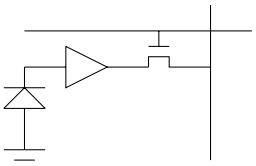
- readout done by transferring one row at a time to the column storage capacitors, then reading out the row one (or more) pixel at a time using the column decoder and multiplexer
- row integration times are *staggerred* by the row/column readout time, 'snap shot' operation can be achieved using a mechanical shutter

CMOS Image Sensor Pixel Architectures

- Passive pixel (PPS)
  - 1 transistor per pixel
  - small pixel, large fill factor, but
  - slow, low SNR

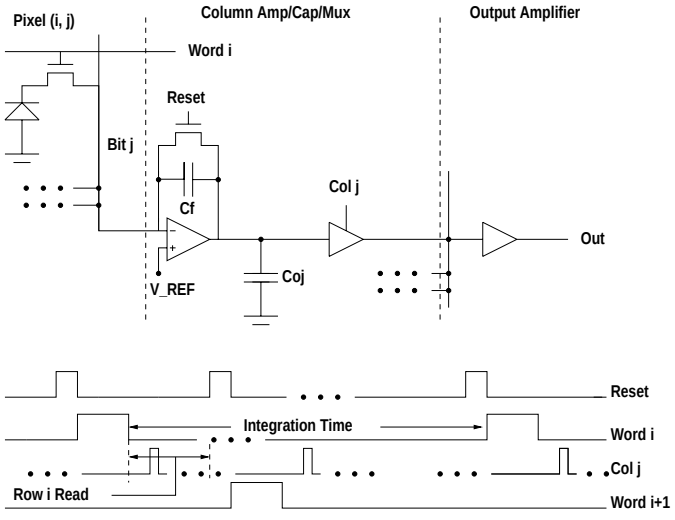


- Active pixel (APS)
  - 3-4 transistors per pixel
  - fast, higher SNR, but
  - larger pixel, lower fill factor



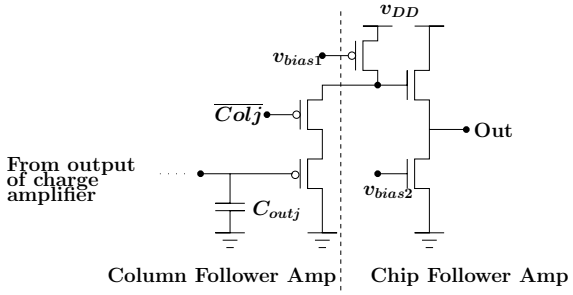
- as technology scaled to 0.5μm, APS pixel size/fill factor no longer a problem
  - current technology of choice

Passive Pixel Sensor (PPS)



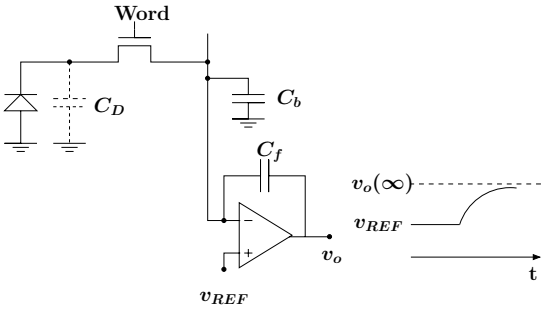
Comments on Operation

- charge is read out via a column charge amplifier
- reading is destructive (much like a DRAM)
- diode reverse bias voltage at end of reading $\approx v_{REF}$
- column and chip amplifiers are simple follower amplifiers



- readout time limited by the time of transferring a row to the output of the charge amplifiers — column readout can be made fast by sizing column and chip follower amplifiers (cannot be ignored though)

PPS Charge to Output Voltage Transfer Function



in steady state, assuming charge Q accumulated on the photodiode at the end of integration (and ignoring "feedthrough" voltage added when the reset transistor is turned off and opamp offset voltage), the output voltage

$$v_o = v_{REF} + \frac{1}{C_f}Q$$

thus the sensor conversion gain is $\frac{q}{C_f}$ (typically reported in μV/electron)
now let's find the sensor voltage swing v_s

the minimum output voltage occurs when $Q = 0$, ignoring dark current we get

$$v_{omin} = v_{REF}$$

the maximum output voltage occurs when the voltage on the diode reaches ground, which gives

$$v_{omax} = v_{REF} + \frac{C_D}{C_f} v_{REF},$$

provided that v_{omax} does not exceed the opamp maximum output voltage v_{Sat} (in this case $v_{omax} = v_{Sat}$)
thus the sensor voltage swing

$$v_s = \min\{\frac{C_D}{C_f} v_{REF}, v_{Sat} - v_{REF}\}$$

Note: since the column amplifier, which can be made quite linear, is used to convert the collected charge to voltage, the output voltage is linear in illumination (F_0), this is similar to CCDs (but different from APS as we shall see)
Special case for examples: we assume $C_f = C_D$, which gives

$$v_s = \min\{v_{REF}, v_{Sat} - v_{REF}\}$$

PPS Readout Speed

- row readout is done in two stages; first the row is transferred to the column capacitors, then the column decoder/multiplexer is used to serially read out the pixel values
- row transfer time can be the dominant readout time
- row transfer time is the time from Word going high to the time v_o is within ϵ of its final value ($v_o(\infty)$)
- for k bits of resolution choose

  $$\epsilon \leq \frac{v_s}{2 \times 2^k} \mathsf{V},$$

  where v_s is the output voltage swing, *e.g.*, for $k = 8$ bits,

  $$\epsilon \leq \frac{v_s}{512}$$

- worst transfer time occurs when $v_o(\infty)$ is maximum, i.e., equal to v_{omax}
- Example: consider a PPS with $n = 256$ rows, $C_D = C_f = 20$fF, and $C_b = n \times 2.6$fF$= 0.6656$pF, find the maximum row transfer time assuming $k = 8$ bits of resolution

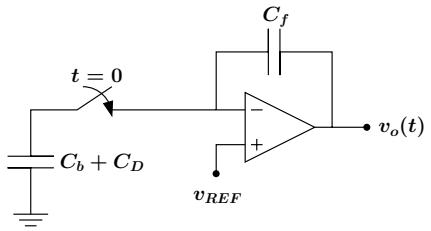
to simplify the analysis we assume:

- single-pole open-loop model for the opamp used in the charge amplifier

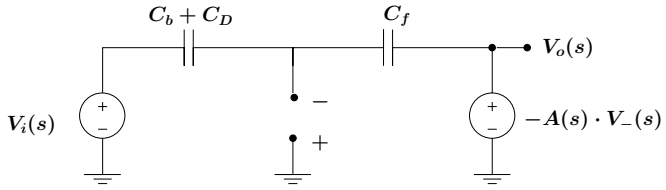
$$\frac{V_o(s)}{V_+(s) - V_-(s)} = A(s) = \frac{A}{1 + (\frac{s}{\omega_o})}$$

  with dc gain $A = 6 \times 10^4$ and 3dB bandwidth $\omega_o = 100$rad/s
- access transistor has negligible ON resistance

to find the row transfer time, we assume that the charge sharing between C_D and C_b occurs instantaneously (the access transistor treated as a short circuit) and use the equivalent circuit



to find the transfer time we substitute the single-pole opamp model to get



the transfer function

$$\frac{V_o(s)}{V_i(s)} \approx -\frac{C_b + C_D}{C_f} \cdot \frac{1}{1 + \frac{s}{(\frac{A\omega_o C_f}{C_b + C_D + C_f})}}$$

thus the time constant

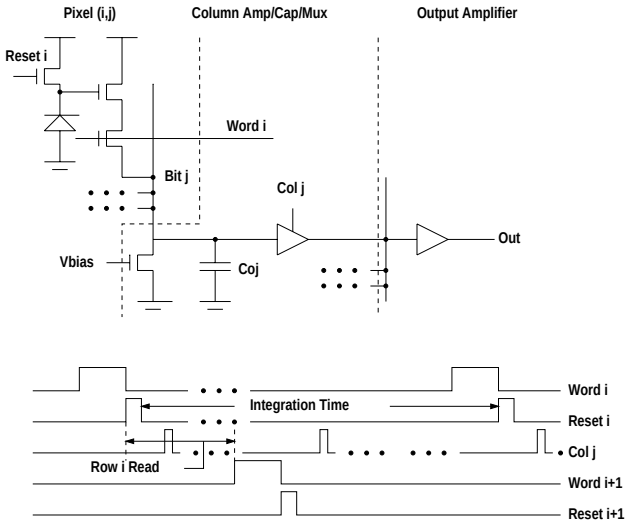
$$\tau = \frac{C_b + C_D + C_f}{A\omega_o C_f} = 5.88\mu\text{s},$$

and

$$\text{worst case row transfer time} = \tau \ln(256 \times 2) = 36.5\mu\text{s}$$

- note that
  - row transfer time increases almost linearly with C_b (and the number of rows)
  - readout time can be reduced by increasing the gain-bandwidth product of the opamp ($A\omega_o$), which would increase power consumption

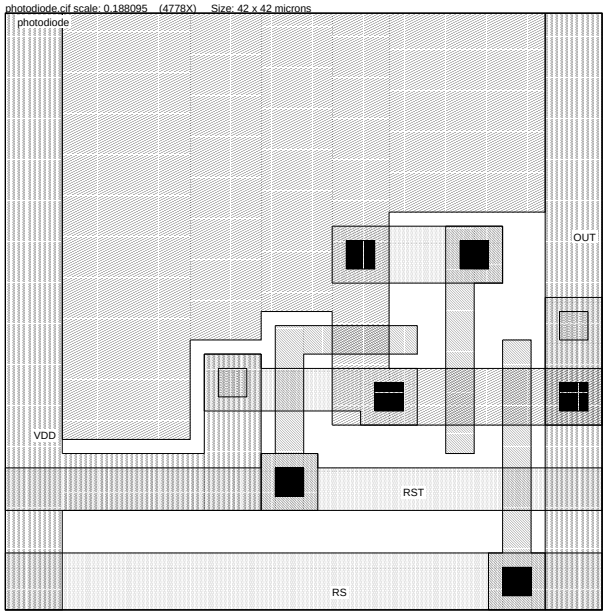
Active Pixel Sensor (APS)



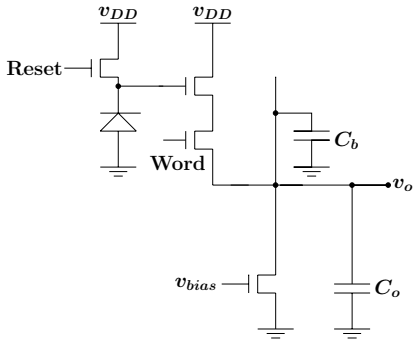
Comments on Operation

- direct integration is used, voltage is read out of the pixel
- output of the photodiode is "buffered" using pixel level follower amplifier — reading is not destructive and can be much faster than PPS
- each row has separate reset (used after reading)
- the photodiode reset voltage $v_D = v_{DD} - v_{TR}$, where v_{TR} is the reset transistor threshold voltage (including body effect)
- by setting the voltage on the reset gate $v_{Reset} \geq v_{TR}$ (instead of ground) during integration, blooming can be controlled (reset transistor doubles as anti-blooming device)
- except for eliminating the charge amplifier, the column amplifier and decoder are identical to PPS

APS Pixel Layout



APS Charge to Output Voltage Transfer Function



in steady state assuming charge Q accumulated on the photodiode at the end of integration and ignoring the voltage drop across the access transistor, the output voltage

$$\begin{aligned} v_o &= v_D - \frac{Q}{C_D} - v_{GSF} \\ &= (v_{DD} - v_{TR}) - \frac{Q}{C_D} - v_{GSF} \end{aligned}$$

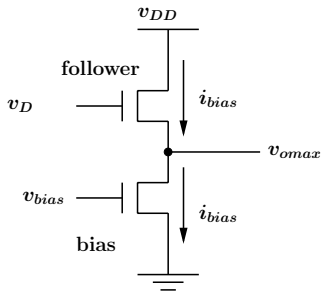
where v_{TR} is the reset transistor threshold voltage (including body effect), and v_{GSF} is the follower transistor gate to source voltage
the sensor conversion gain is thus $\frac{q}{C_D}\mu$V/electron
to keep the bias transistor in saturation we choose

$$v_{omin} = v_{bias} - v_{TB}$$

where v_{TB} is the bias transistor threshold voltage
the maximum output voltage occurs when $Q = 0$, thus

$$v_{omax} = v_{DD} - v_{TR} - v_{GSF}$$

to find v_{GSF} consider the circuit (in steady state)



where i_{bias} is the column amplifier bias current
assuming the static first order MOS transistor model, we get

$$i_{bias} = \frac{k_n}{2} \cdot \frac{W_F}{L_F}(v_{GSF} - v_{TF})^2$$

where W_F and L_F are the follower transistor width and length, and v_{TF} is its threshold voltage
thus we get that

$$v_{GSF} = v_{TF} + \sqrt{\frac{2L_F}{k_n W_F} i_{bias}}$$

thus the voltage swing is given by

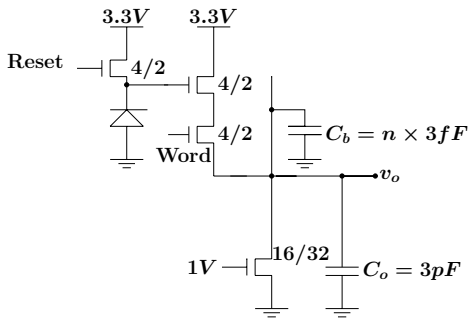
$$v_s = v_{DD} - v_{TR} - v_{GSF} - v_{bias} + v_{TB}$$

Remarks:

- the available well capacity $Q_{max} = v_D \times C_D$ cannot be fully utilized, since v_{omin} is achieved before the diode voltage drops to ground
- since the collected charge is converted to voltage using the diode capacitance C_D, the output voltage can be somewhat nonlinear in illumination (F_0)

APS Readout Speed

- readout time for a row is the sum of the time to transfer the row to the column capacitors and the time to read out the pixel values via the column multiplexer (column readout time)
- unlike for PPS, APS column readout time (and not row transfer time) is the real performance limiter
- Example: consider an APS implemented in the 0.5μ CMOS technology described in Handout 2 with $n = 256$ rows, $v_{TF} = 0.9$V, $v_{TR} = 1.1$V, $v_{TB} = 0.8$V, and the parameters shown in the figure



find the row transfer time assuming $k = 8$ bits of resolution

first let's compute the output voltage swing v_s

the minimum output voltage

$$v_{omin} = v_{bias} - v_{TB} = 0.2\text{V}$$

the bias current using $k_n = 188\mu\text{A/V}^2$

$$i_{bias} = 1.88\mu\text{A}$$

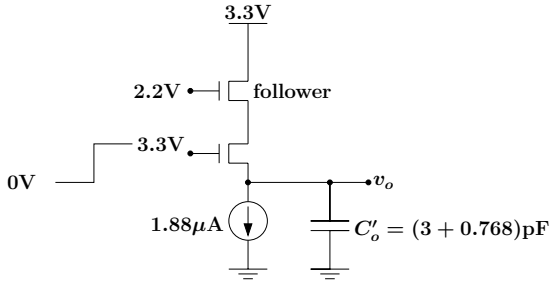
thus the maximum output voltage

$$v_{omax} = v_{DD} - v_{TR} - v_{GSF} = 3.3 - 1.1 - 1.0 = 1.2\text{V},$$

and

$$v_s = 1.0\text{V}$$

to find the worst case row transfer time, we use the equivalent circuit



from KCL

$$1.88 \times 10^{-6} + C'_o \cdot \frac{dv_o}{dt} = 188 \times 10^{-6} \times (1.3 - v_o)^2$$

now we solve the differential equation to find the row transfer time t_{row}

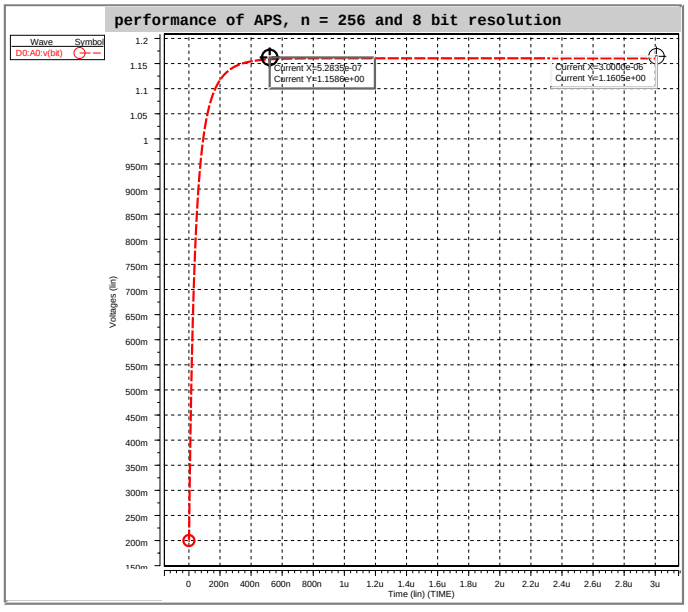
$$\int_0^{t_{row}} dt \;=\; t_{row} = C'_o \int_{0.2}^{1.198} \frac{dv_o \times 10^6}{188(1.3 - v_o)^2 - 1.88}$$

$$
\begin{aligned}
&= \frac{10^6 \times C'_o}{188} \int_{0.2}^{1.198} \frac{dv_o}{(1.3 - v_o)^2 - 10^{-2}} \\
&= \frac{10^6 \times C'_o}{188 \times 0.2} \int_{0.2}^{1.198} \left(\frac{1}{1.2 - v_o} - \frac{1}{1.4 - v_o}\right) dv_o \\
&= \frac{3.768 \times 10^{-6}}{188 \times 0.2} \times 4.43 \\
&= 444\text{ns}
\end{aligned}
$$

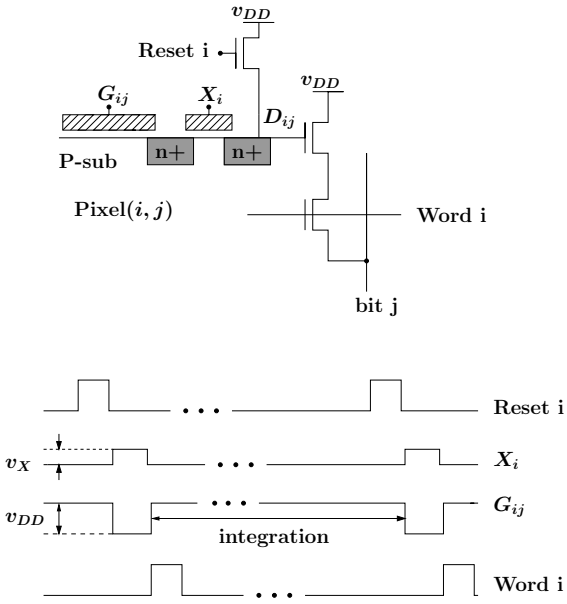
where the lower limit of the integral is v_{omin} and the upper limit is $(v_{omax} - \frac{v_s}{512})$

- note that
  - the row transfer time increases linearly in C_b (same as PPS)
  - it is considerably faster than PPS (our estimate is, however, somewhat optimistic as you can see from the HSPICE simulations)
  - PPS can be made as fast using a fast charge amplifier, but power consumption would be much higher than APS

APS Readout Speed from HSPICE Simulation



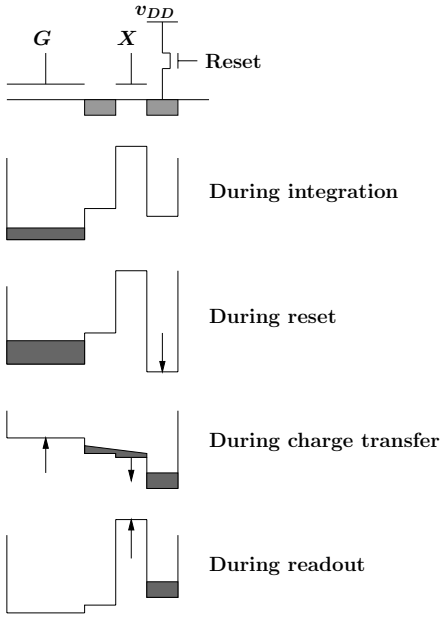
CMOS Photogate APS



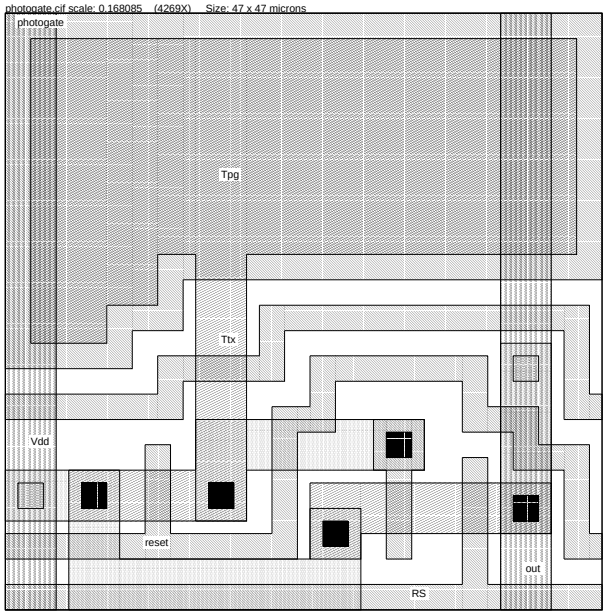
Comments on Operation

- column and chip circuits are identical to photodiode APS
- before reading a row
  - floating node D_{ij} is reset to $v_D = v_{DD} - v_{TR}$
  - to transfer the accumulated charge on the photogate to the floating node D_{ij}, the transfer gate is turned to an intermediate voltage $\leq \frac{v_{DD}}{2}$V and the gate voltage is lowered to 0V (CCD like operation)
- the transfer gate can also be kept at a constant intermediate voltage throughout to avoid blooming and to reduce switching noise
- well capacity is determined by the voltage swing on the floating node and its capacitance C_d
- rest of operation identical to photodiode APS
- the transfer gate, reset transistor and follower circuit is identical to the structure of the CCD output amplifier

Potential Well Diagram for Photogate APS



Photogate APS Pixel Layout



Advantages and Disadvantages of Photogate

- advantages
  - conversion gain $\frac{q}{C_d}$ is independent of detector, can achieve higher conversion gain than photodiode APS (lower QE, however)
  - C_d very useful when performing $correlated\ double\ sampling$ (CDS) (as we will see later) and may be used to perform frame differencing (read-reset-read new frame)
- disadvantages
  - more devices (larger pixel or lower fill factor than photodiode)
  - poor blue response