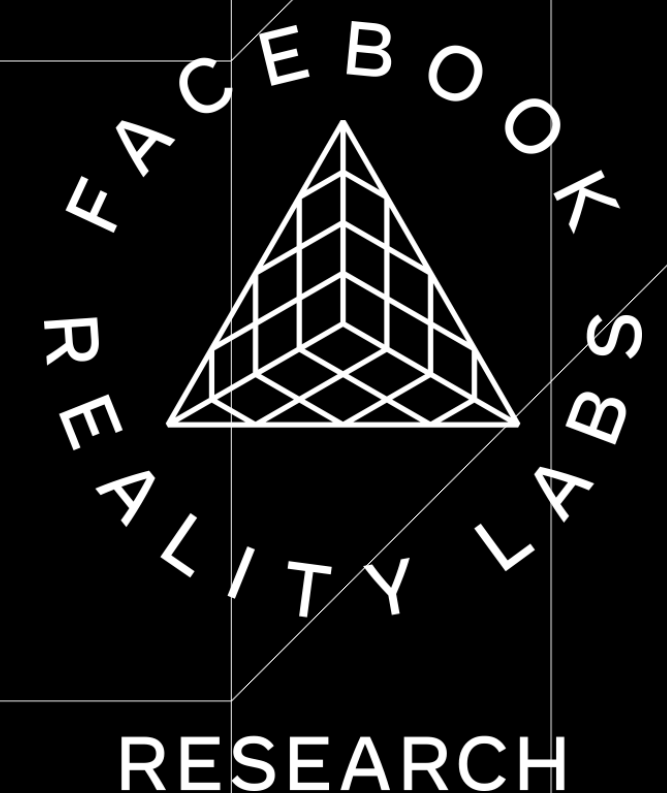


# Sensing and Computing Technologies for AR/VR

**Chiao Liu**

Director of Research  
Facebook Reality Labs

March 15, 2021



# Quick Bio

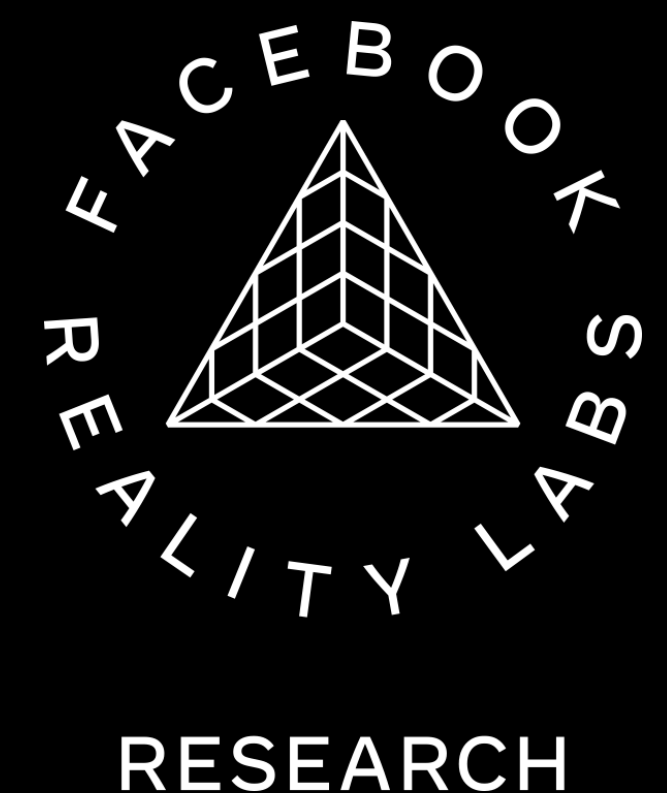
PhD, EE, Stanford University

Senior Scientist, Canesta Inc (now part of Microsoft), focusing on 3D ToF sensors

Technical Fellow, Fairchild Imaging, focusing on scientific and medical imaging

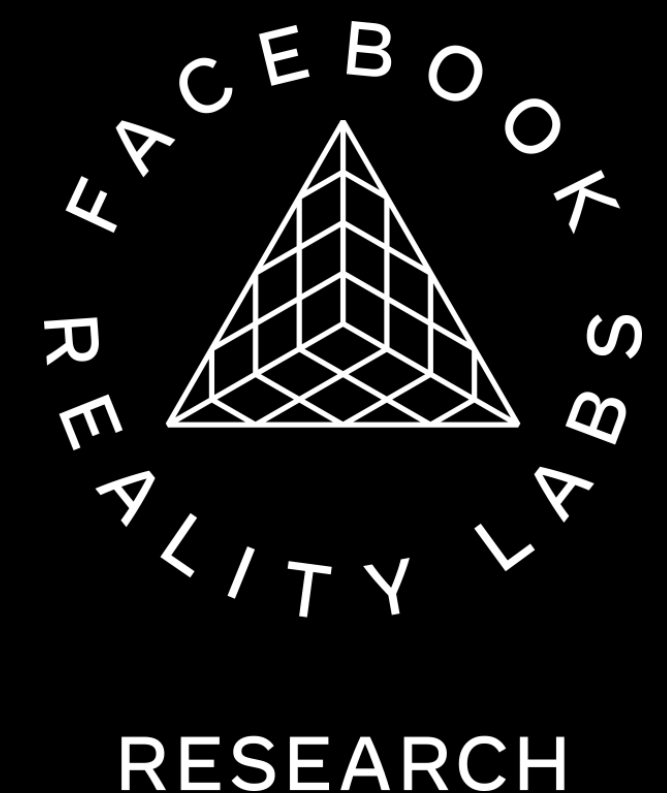
Principle Architect, Microsoft, part of 1<sup>st</sup> gen HoloLens AR team

Director of Research, Facebook Reality Labs, leading Sensors and Systems Research team



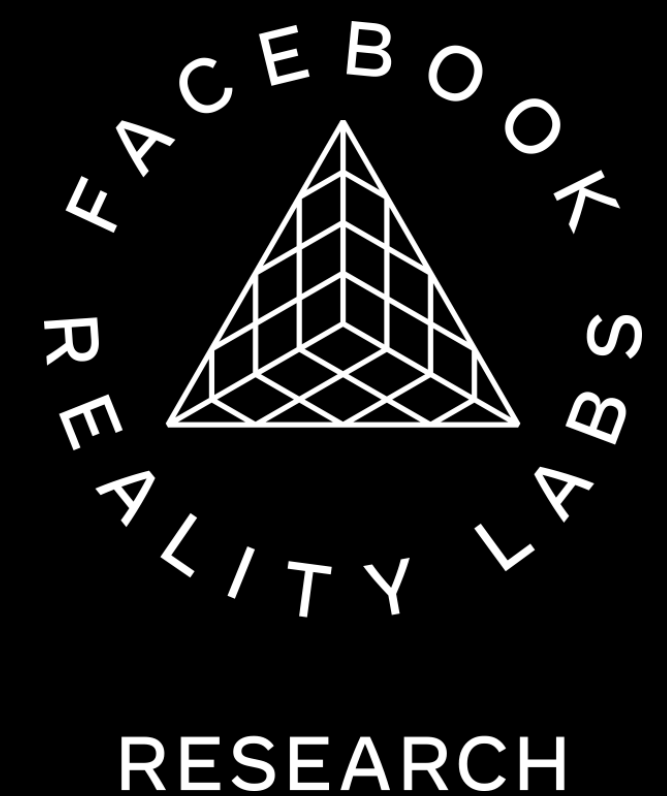
# Agenda

1. Image Sensing in AR/VR
2. CMOS Image Sensor Basics
3. A Digital Pixel Sensor w/ Ultra Wide Dynamic Range
4. A Distributed Sensing and Compute Architecture

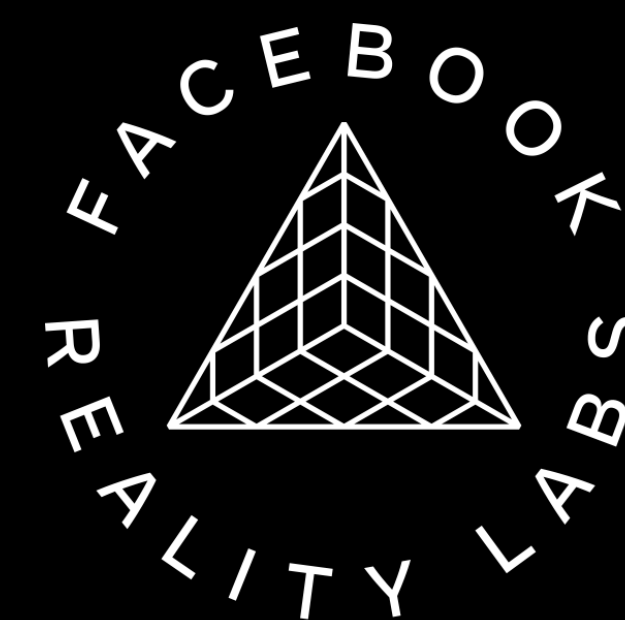
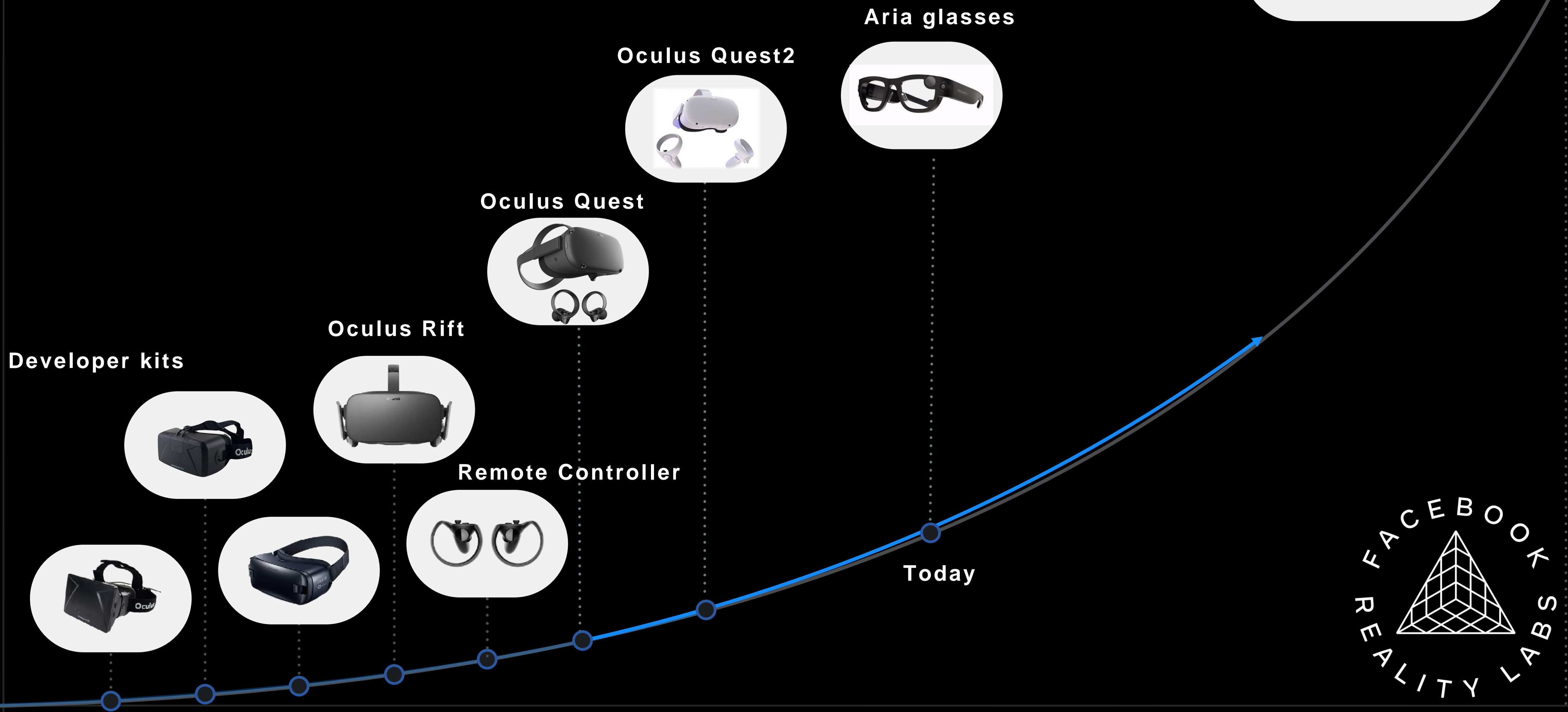


# Agenda

1. Image Sensing in AR/VR
2. CMOS Image Sensor Basics
3. A Digital Pixel Sensor w/ Ultra Wide Dynamic Range
4. A Distributed Sensing and Compute Architecture



# AR/VR: the next generation computing platform



RESEARCH

 **oculus quest**





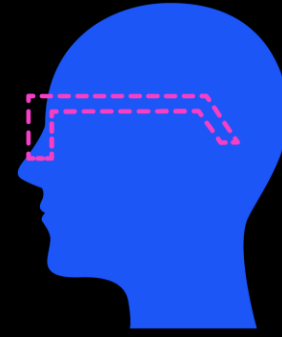
Quest 2

# The AR Glasses





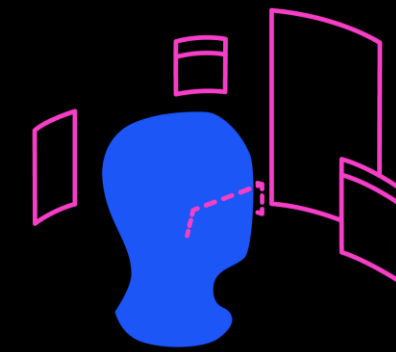




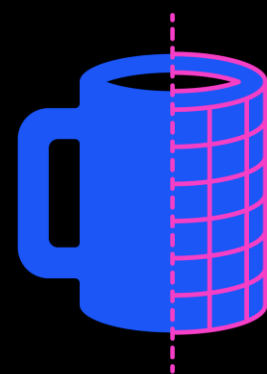
**SYSTEM  
DESIGN**



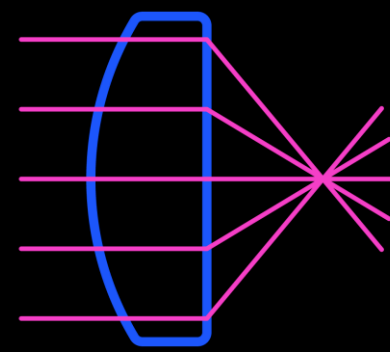
**INTERACTION**



**UX**



**SENSING AND  
TRACKING**

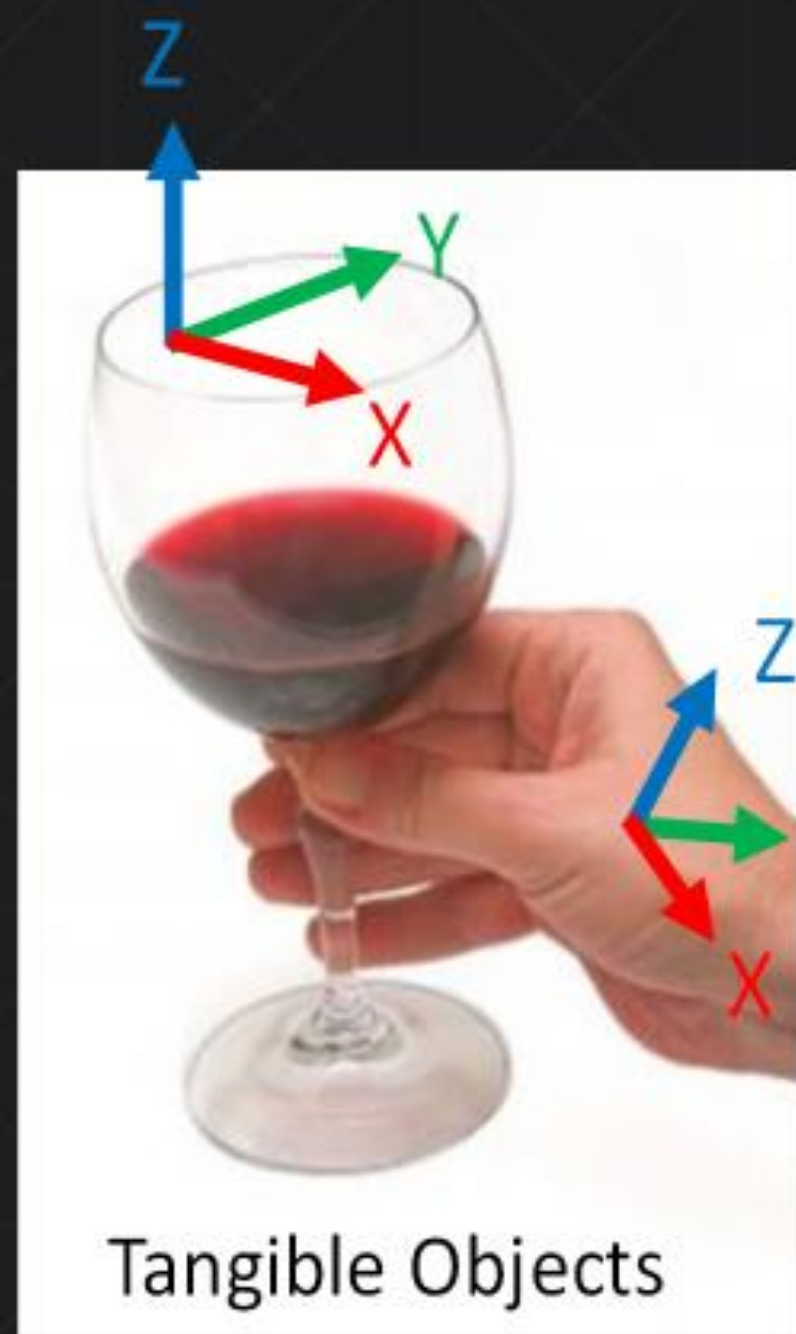


**OPTICS &  
DISPLAYS**

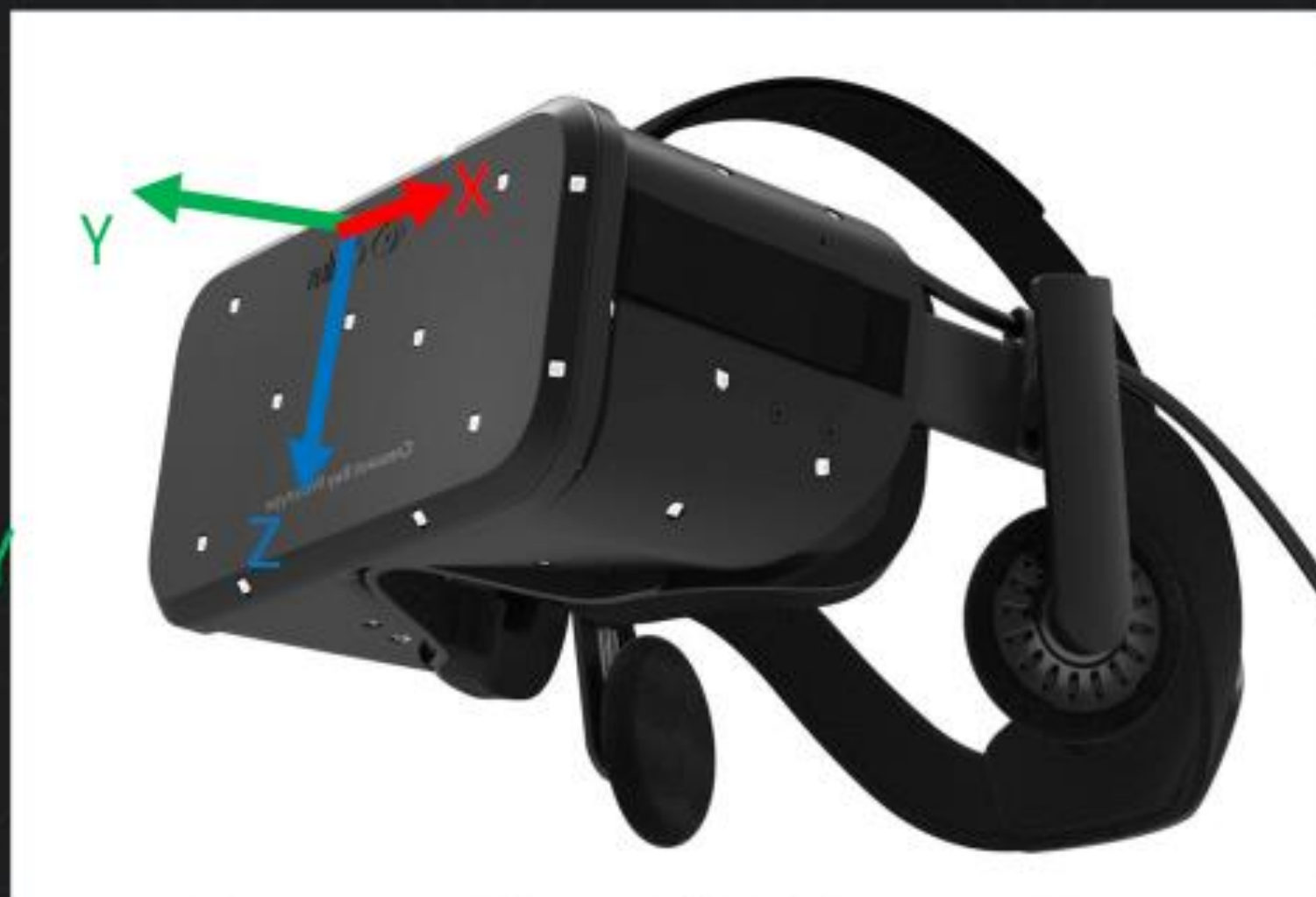


**AUDIO**

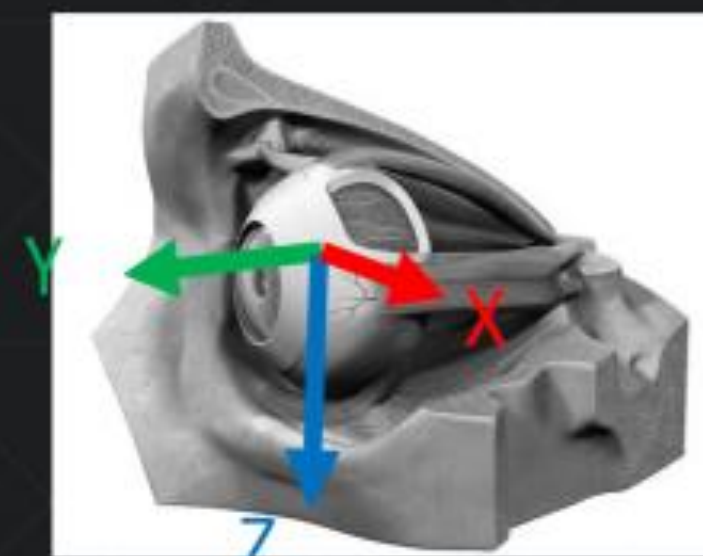
# / Core Sensing and Tracking Functions



Tangible Objects  
& Hands



Head/Eye-box



Eye

World frame → Objects and hand tracking → Head tracking → Eye tracking

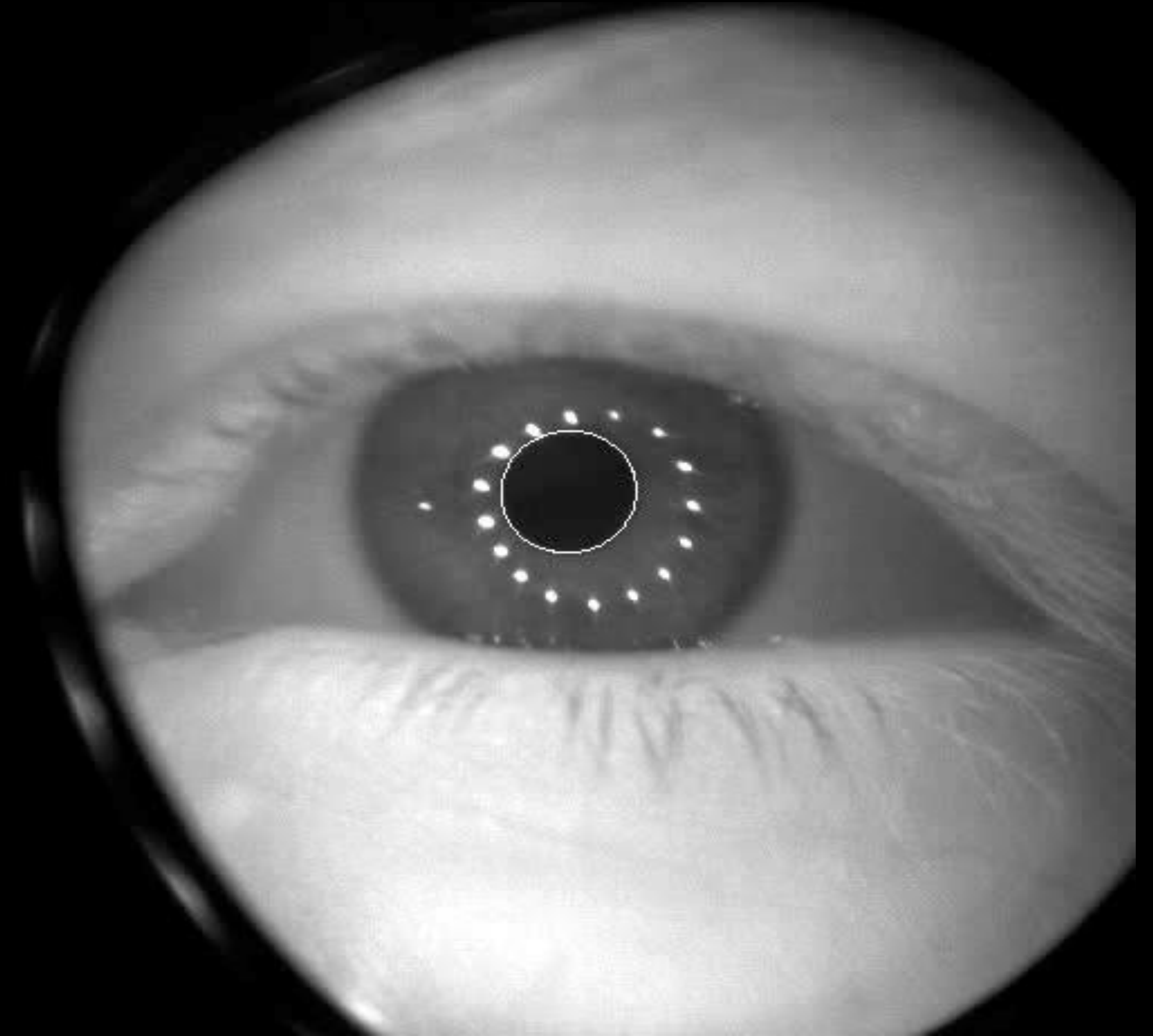
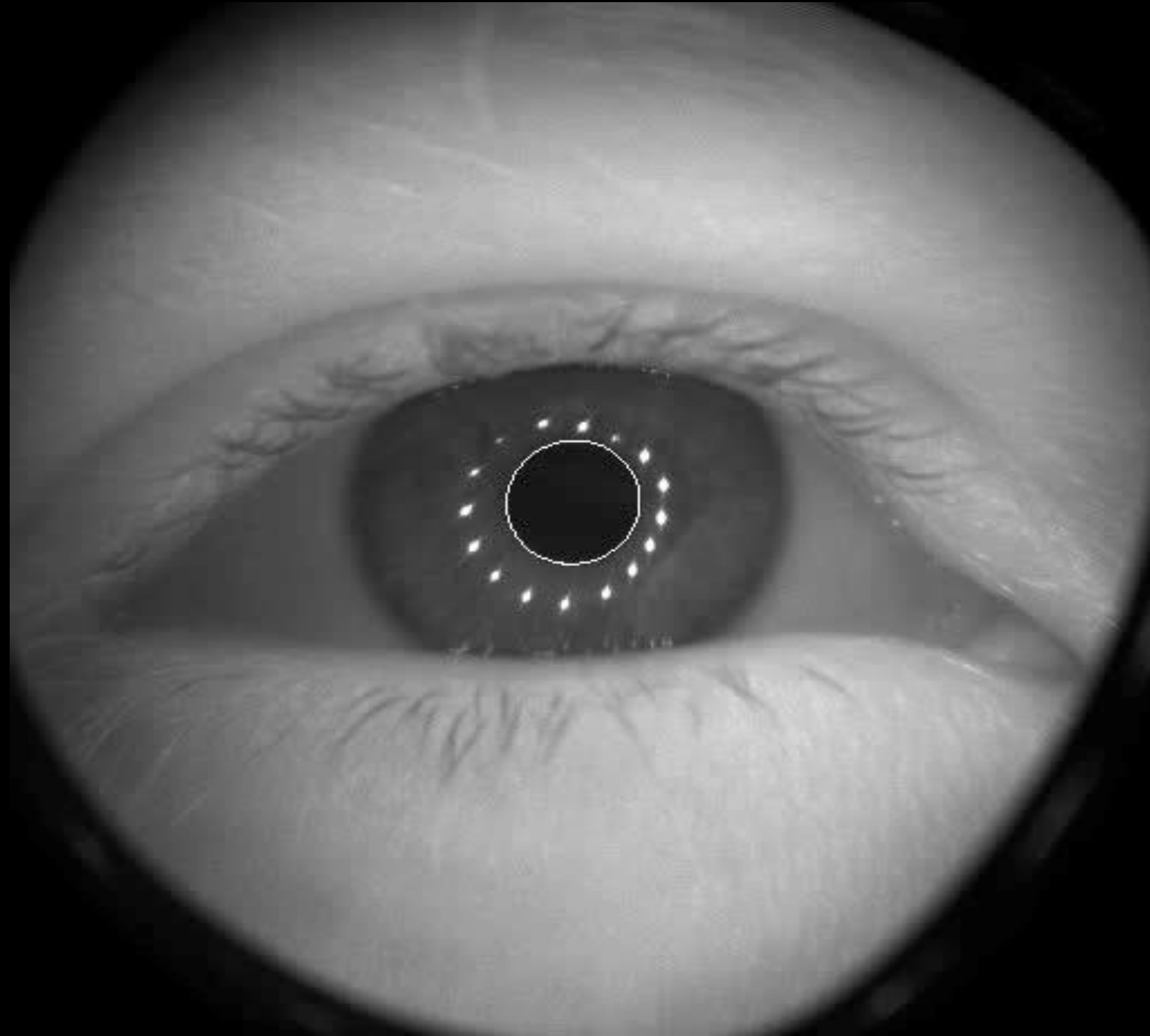
# Sensor Use Case --- SLAM



# Sensor Use Case --- Hand Tracking

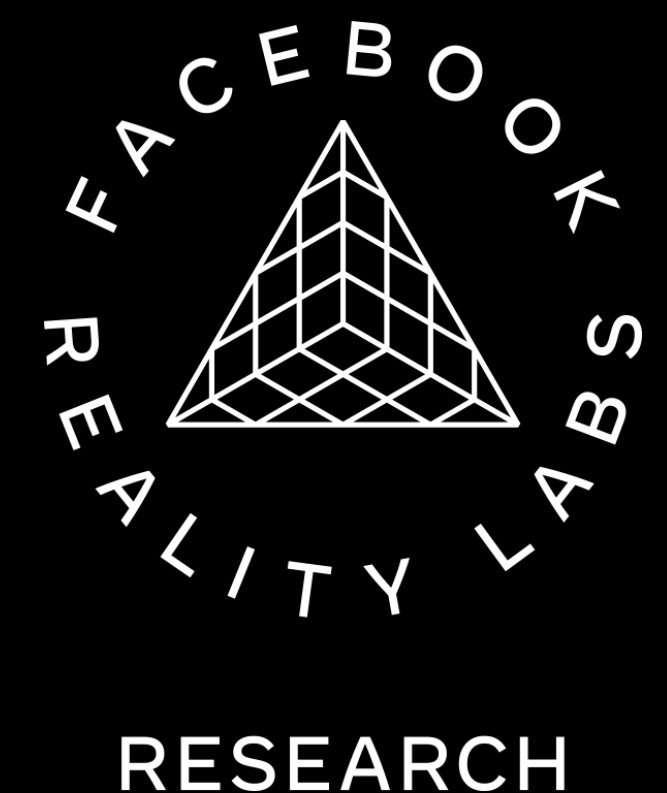


# Sensor Use Case --- Eye Tracking



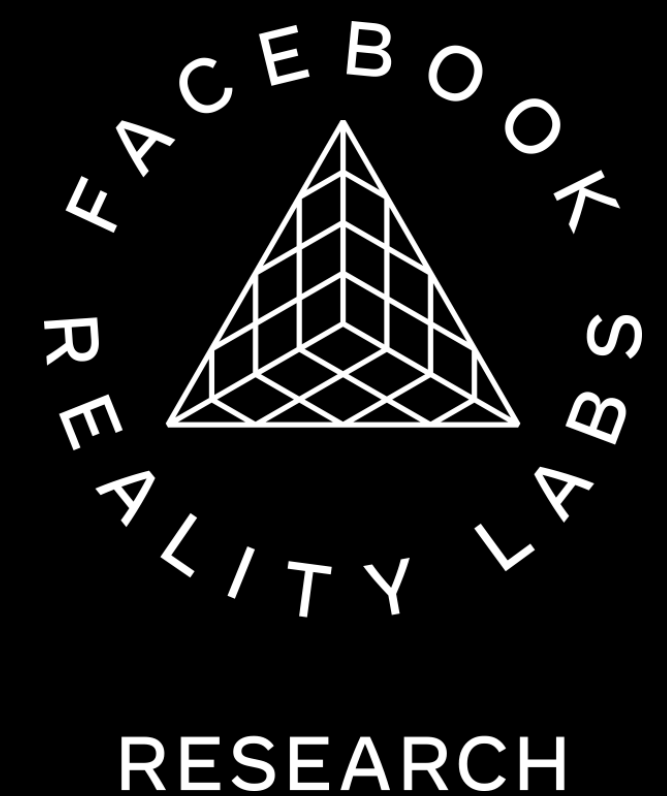
# Sensor and Camera Requirements

- Global shutter sensors, ideal for computer vision applications
- Low light sensitivity, high VIS & NIR sensitivity
- Wide dynamic range
- High frame rate, short exposure time for minimal motion blur
- Ultra-low power consumption for always-on functions
- Small form factor, low weight



# Agenda

1. Image Sensing in AR/VR
2. CMOS Image Sensor Basics
3. A Digital Pixel Sensor w/ Ultra Wide Dynamic Range
4. A Distributed Sensing and Compute Architecture





# Properties of Light

- Visible and near infrared photon energy:

Violet: 400 nm ( $E_{ph} = 3.1$  eV)  
Blue: 450 nm ( $E_{ph} = 2.76$  eV)  
Cyan: 500 nm ( $E_{ph} = 2.48$  eV)  
Green: 550 nm ( $E_{ph} = 2.27$  eV)  
Yellow: 600 nm ( $E_{ph} = 2.08$  eV)  
Red: 700 nm ( $E_{ph} = 1.77$  eV)  
Infrared:  $> 800$  nm ( $E_{ph} < 1.55$  eV)

- Illuminance and Photon Flux (photons/cm<sup>2</sup>.sec)

bright sunlight  $\approx 10^5 - 10^6$  Lux

clear sky  $\approx 10^4$  Lux, or  $F_0 = 10^{17}$

room light  $\approx 10$  Lux, or  $F_0 = 10^{13}$

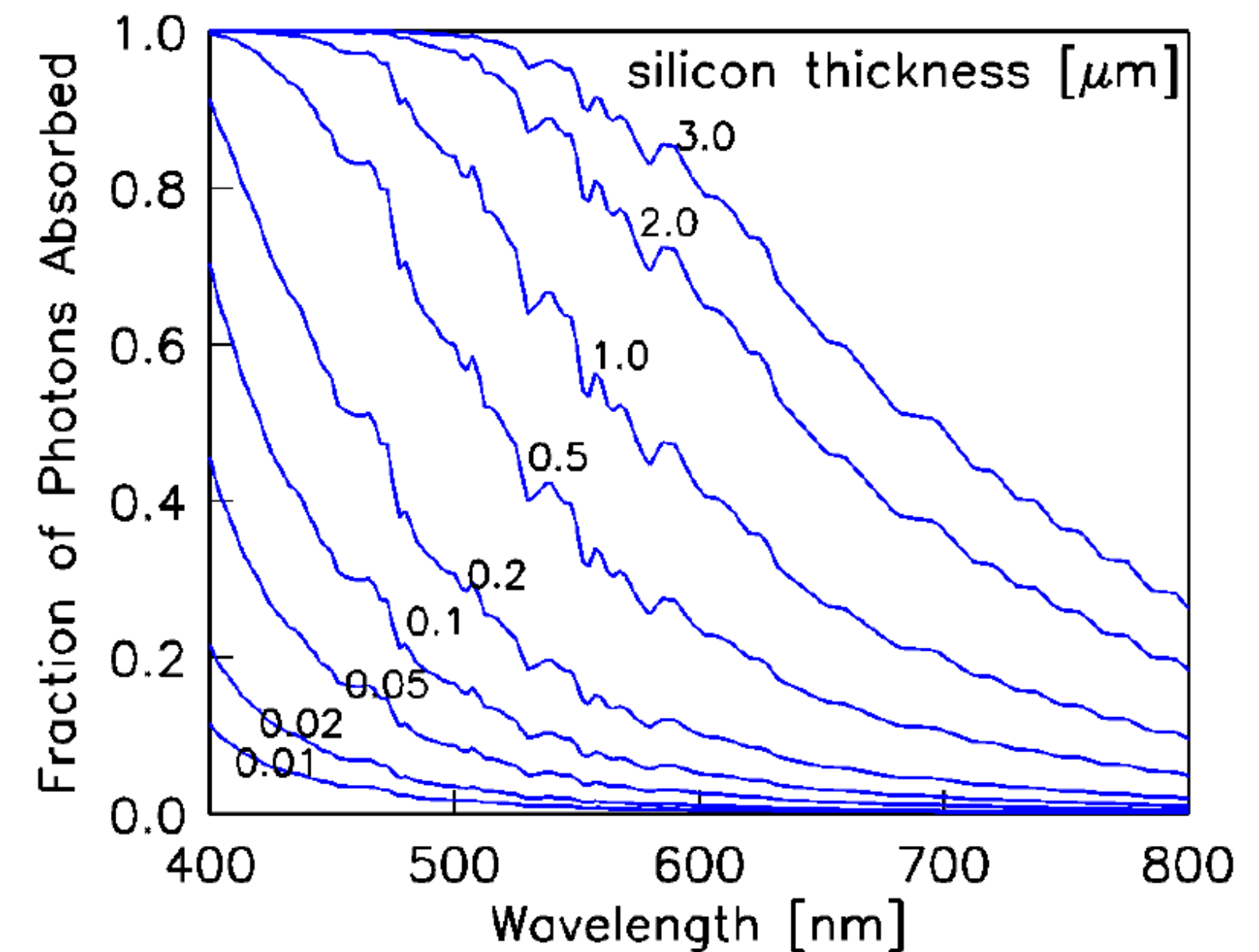
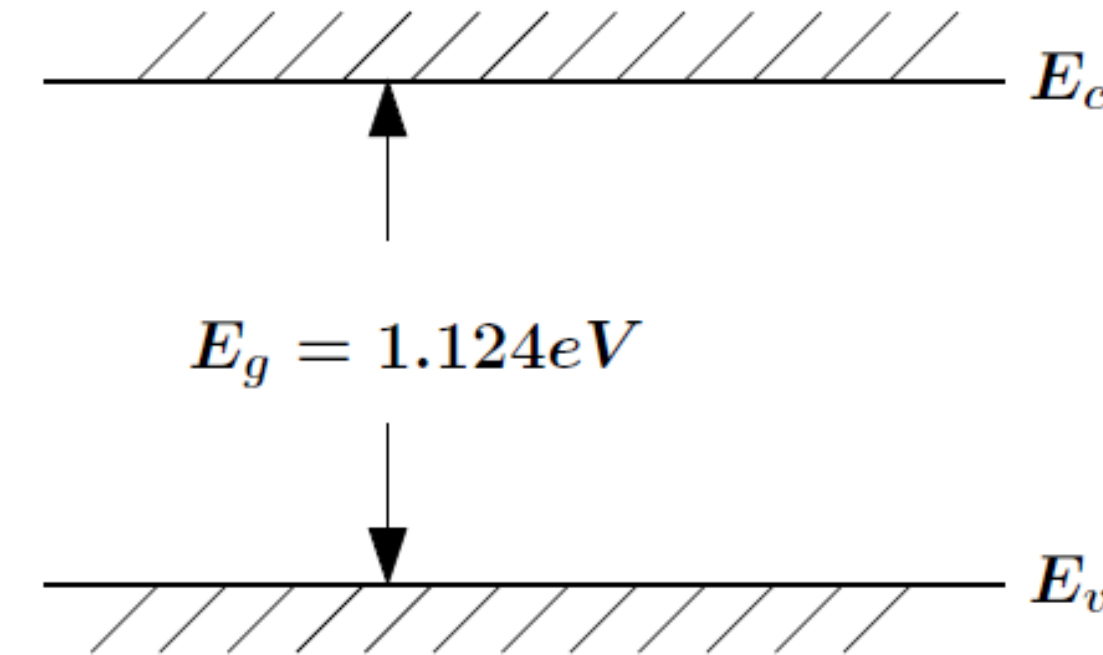
full moon  $\approx 0.1$  Lux, or  $F_0 = 10^{11}$

moonless night  $\approx 10^{-4}$  Lux, or  $F_0 = 10^8$

- Scene dynamic range:
  - $10^6:1 \Rightarrow 120$ dB
  - $10^8:1 \Rightarrow 160$ dB

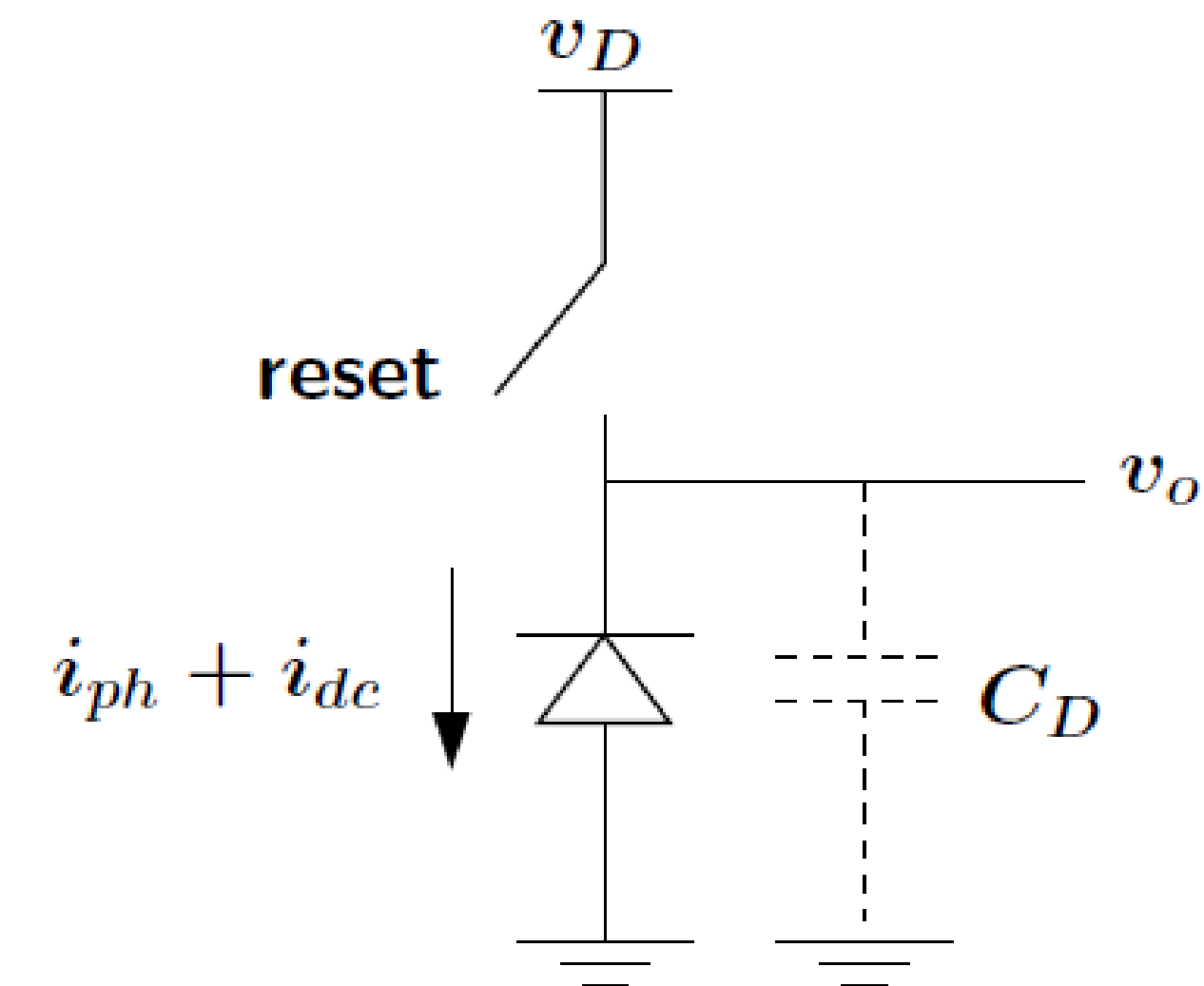
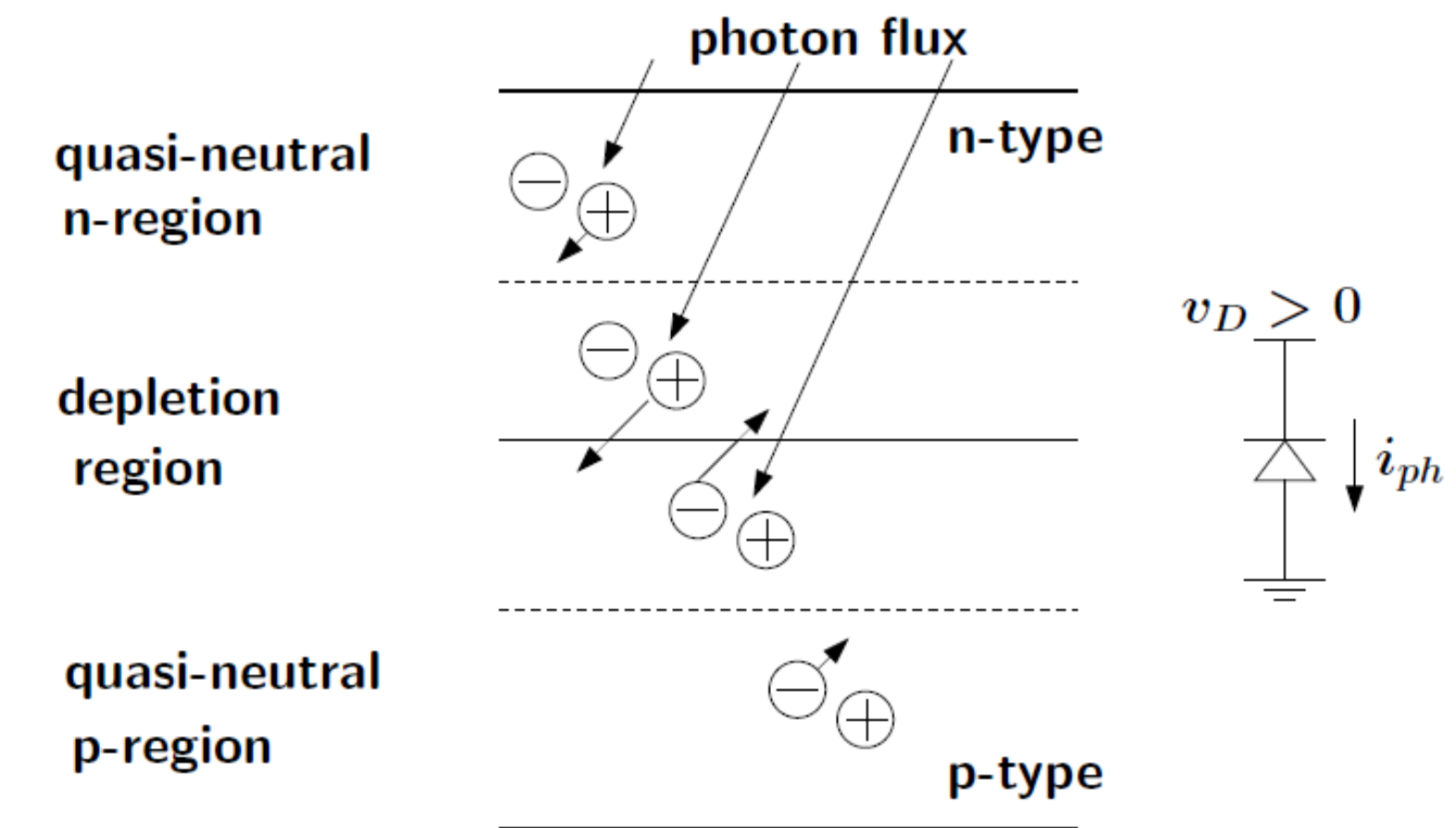
# Light Absorption in Silicon

- Energy band diagram of silicon
- Coincidentally (and luckily) photons in visible range have enough energy to generate electron-hole pairs in silicon
  - Enables us to leverage the technology sophistication, scale and cost benefits of semiconductor industry
- Photon absorption efficiency is a function of wavelength and silicon thickness

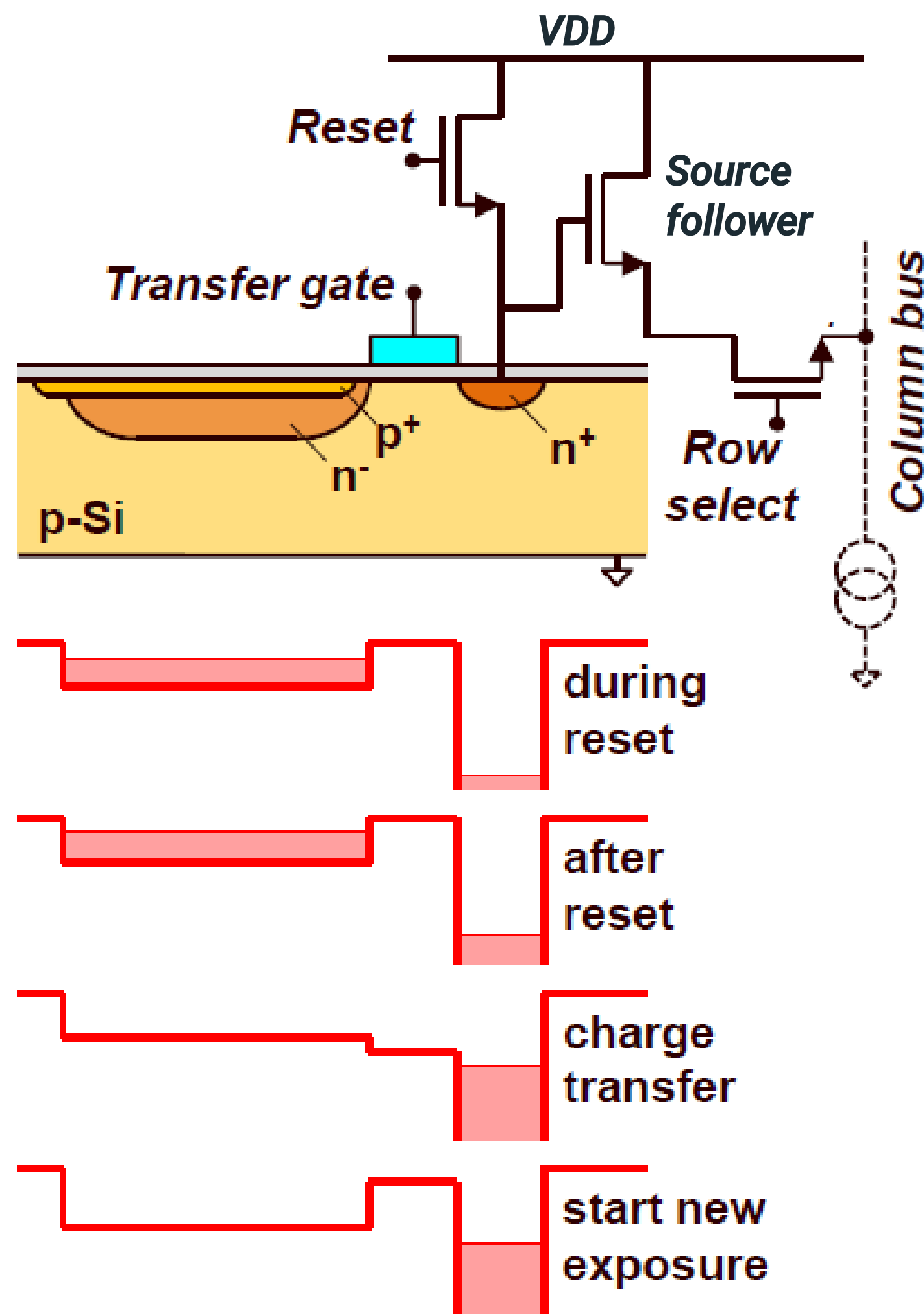


# Signal Generation in Pixel

- Silicon converts photons to electron-hole pairs
- Photodiode accumulates photoelectrons during exposure time
- Photoelectrons convert to voltage signal
  - Either directly on the diode by its parasitic capacitor
  - Or by a separate capacitor where charge is transferred from the photodiode after exposure time
- Voltage signal is quantized to digital bits by an ADC
  - One ADC per sensor
  - One ADC per column
  - One ADC per pixel

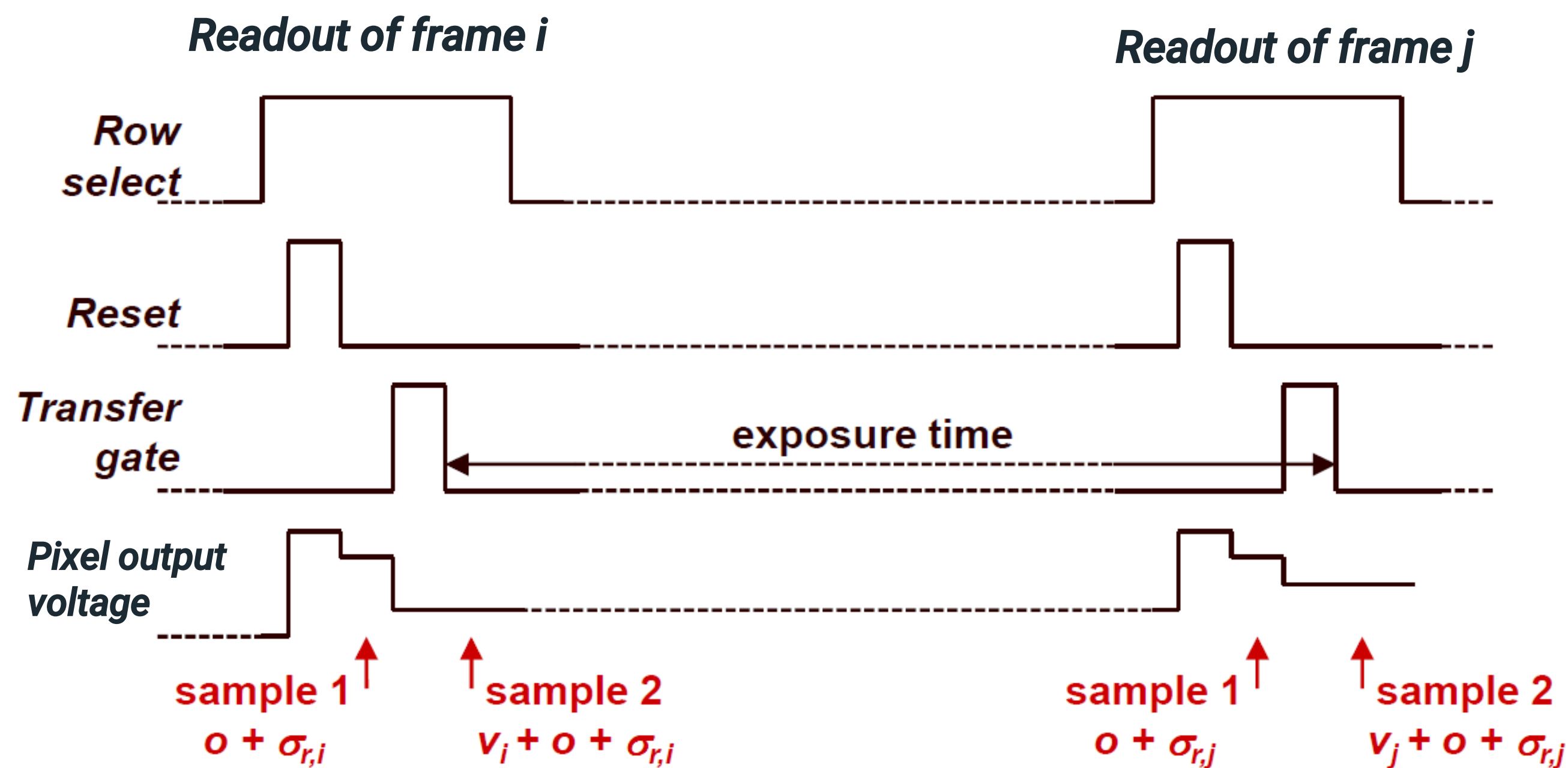


# A 4-Transistor Rolling Shutter Pixel



- N- layer is fully depleted and isolated when transfer gate is off. Photoelectrons stored in N- during exposure
- Transfer gate serves as voltage controlled charge flow barrier
- Reset pulse set N+ (floating diffusion node) to a fixed voltage
- Charge is *fully transferred* from N- to N+ after transfer gate pulse
- Source follower gate serves as buffer/amplifier, driving column bus when Row select gate is on
- Perform correlated double sampling (CDS) by read the column bus voltage twice (more on next slide)
- Rolling Shutter operation

# Correlated Double Sampling (CDS)



After CDS (Correlated Double Sampling):

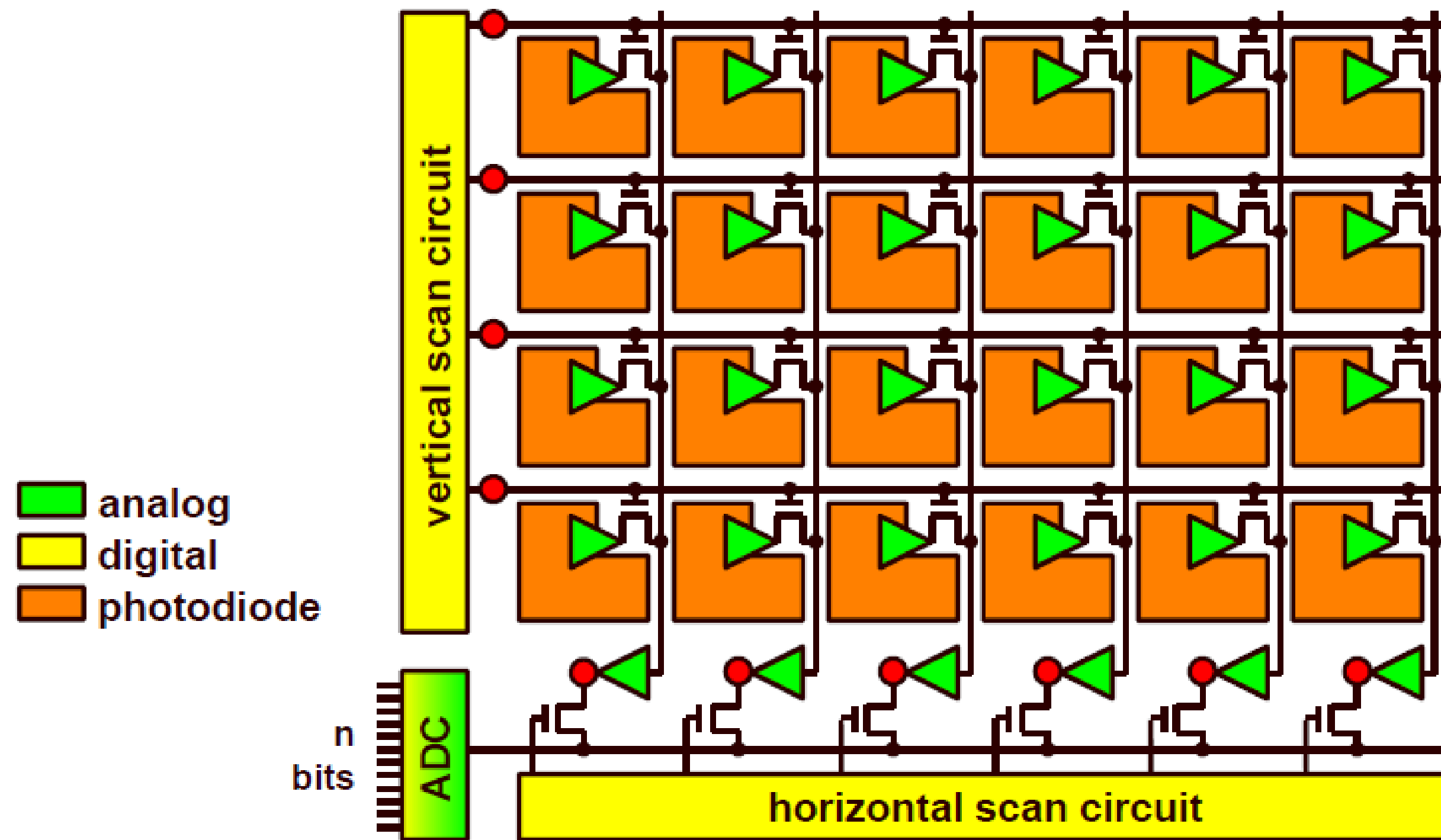
$$v_i + o + \sigma_{r,i} - (o + \sigma_{r,i}) = v_i$$

$$v_j + o + \sigma_{r,j} - (o + \sigma_{r,j}) = v_j$$

$i, j, k, \dots$  : frame number,  
 $v_i$  : video signal after cycle *i* (V),  
 $o$  : pixel offset (V)  
 $\sigma_{r,i}$  : reset or kTC noise of cycle *i* (V).

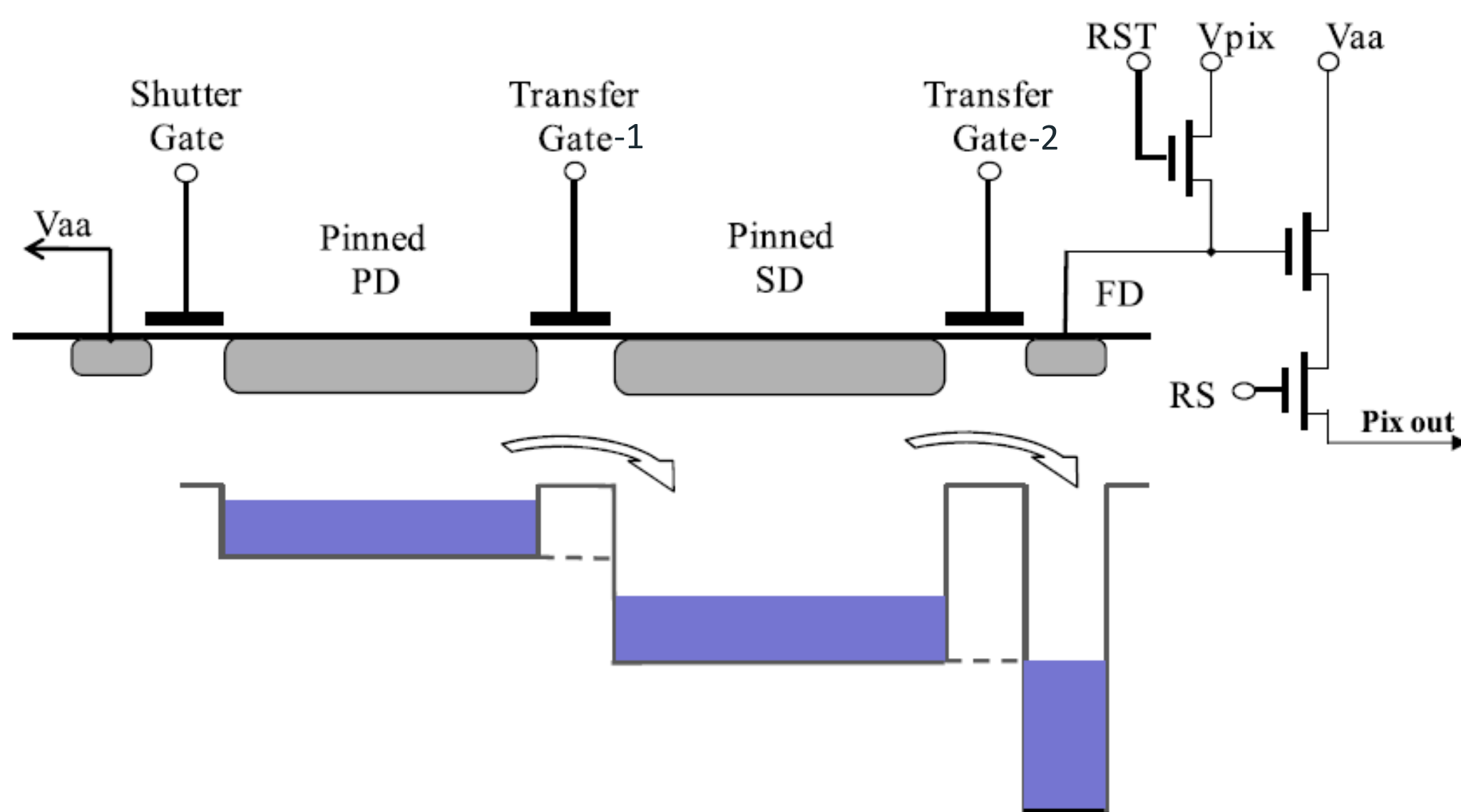
- Multiple noise sources
  - Reset noise --- KTC
  - Thermal and 1/f noise
  - Fixed pattern noise (pixel offset)
- By reading two correlated samples --- one after reset (sample 1) and one after charge transfer (sample 2) --- and subtract, the KTC and FPN noise can be cancelled

# Image Sensor Architecture



- Array of pixels
  - Photodiode
  - Amplifier
  - Selection switch
- During readout, vertical scan circuit select one row at a time
- Horizontal scan circuit multiplex column bus
- Pixel voltage quantized to digital pixel value
  - One ADC for the whole array
  - Or one ADC per column for higher frame rate

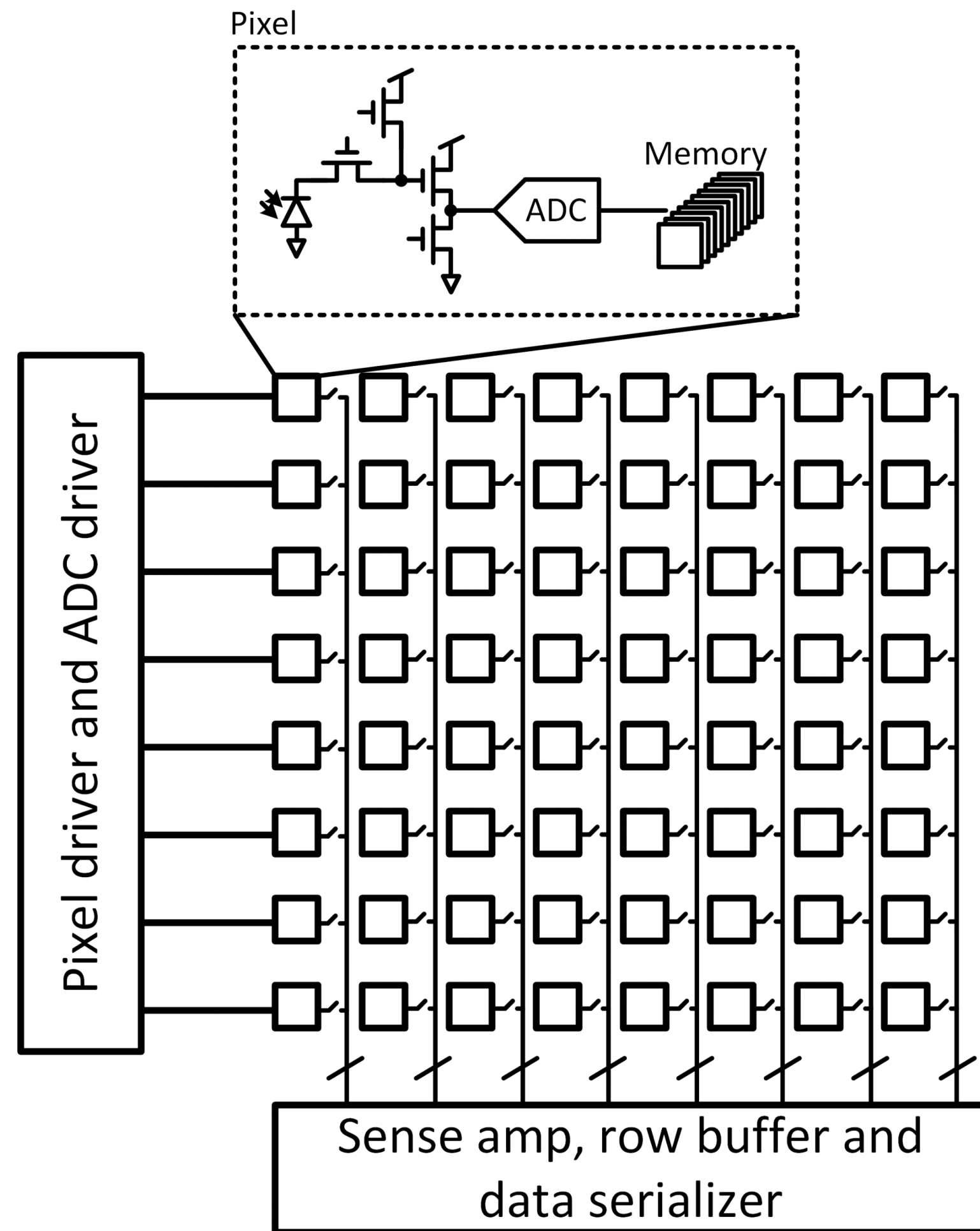
# Global Shutter Pixel



- 6 transistors per pixel
- Shutter Gate serves as global electronic shutter
- All pixels start exposure at same time by turning off Shutter Gate
- All pixels finish exposure at same time by turning on Transfer Gate-1 and move charge from Pinned PD to Pinned SD for storage
- A light shield (typically metal layer) on top of Pinned SD is required to prevent light leakage
- CDS is done by
  - Reset and read FD node (sample 1)
  - Turn on Transfer Gate-2, move charge from Pinned SD to FD, read again (sample 2)
- Larger pixel size and more challenging than rolling shutter pixel

S. Velichko, et. al., "CMOS Global Shutter Charge Storage Pixels With Improved Performance",  
IEEE Trans. on Electron Devices, January 2016

# Digital Pixel Sensor (DPS)

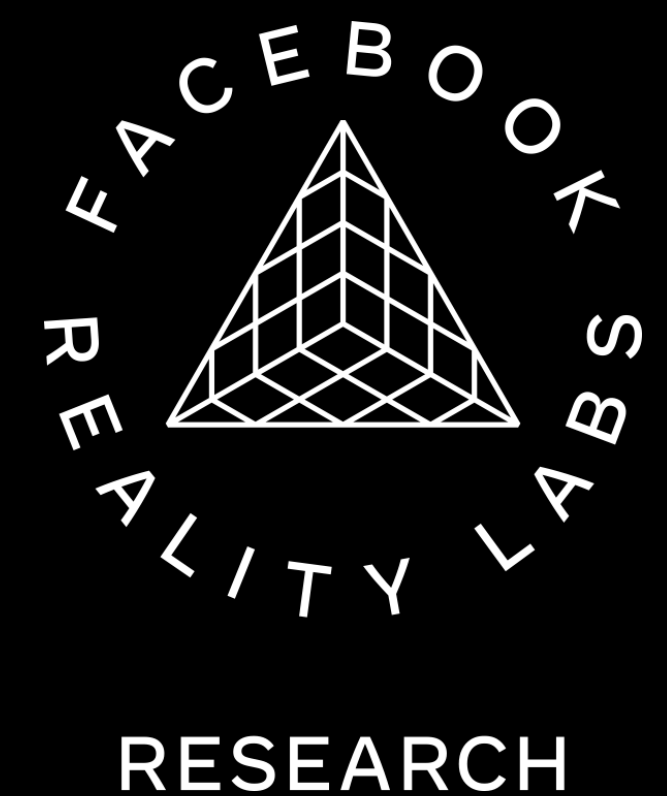


- Each pixel contains an ADC and digital memory
- Pixel data is stored in digital memory and readout as digital bits
- ADC circuit has to be small enough to keep pixel size reasonable
- Now you can treat the pixel array as your frame buffer memory!
- DPS opens new possibilities, we will demonstrate one example in next section

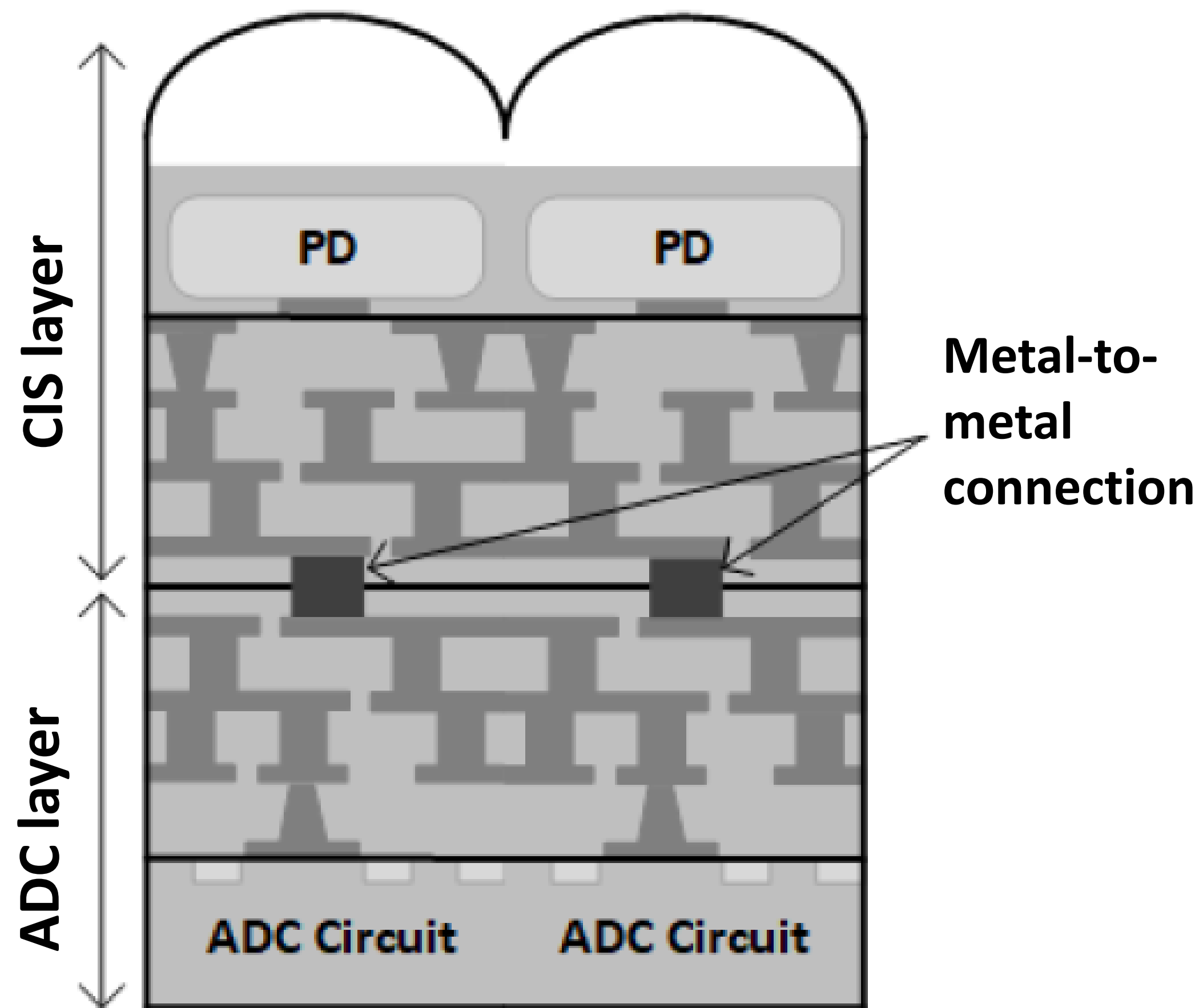


# Agenda

1. Image Sensing in AR/VR
2. CMOS Image Sensor Basics
3. A Digital Pixel Sensor w/ Ultra Wide Dynamic Range (IEDM 2020)
4. A Distributed Sensing and Compute Architecture

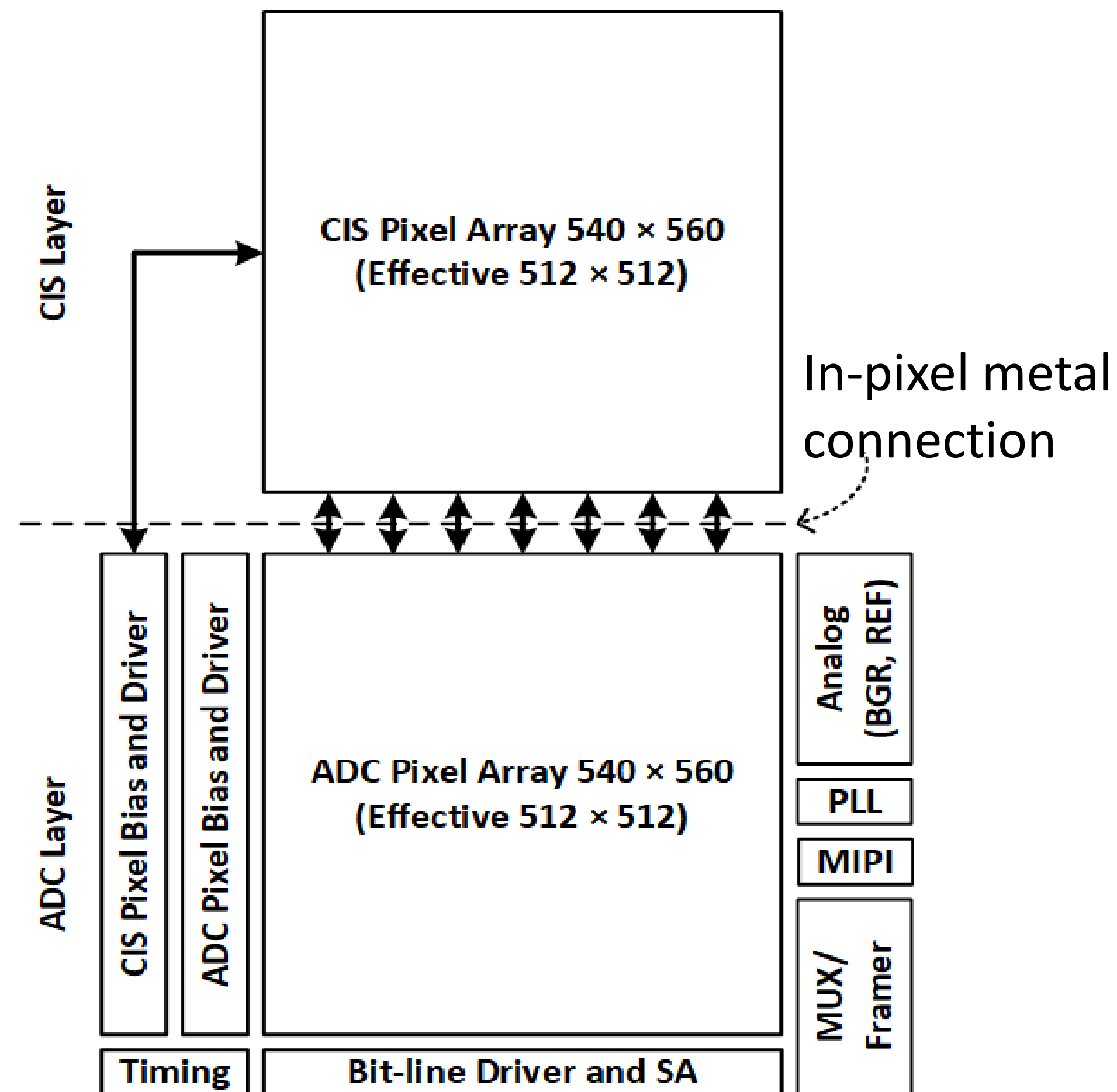


# DPS in a Stacked Process

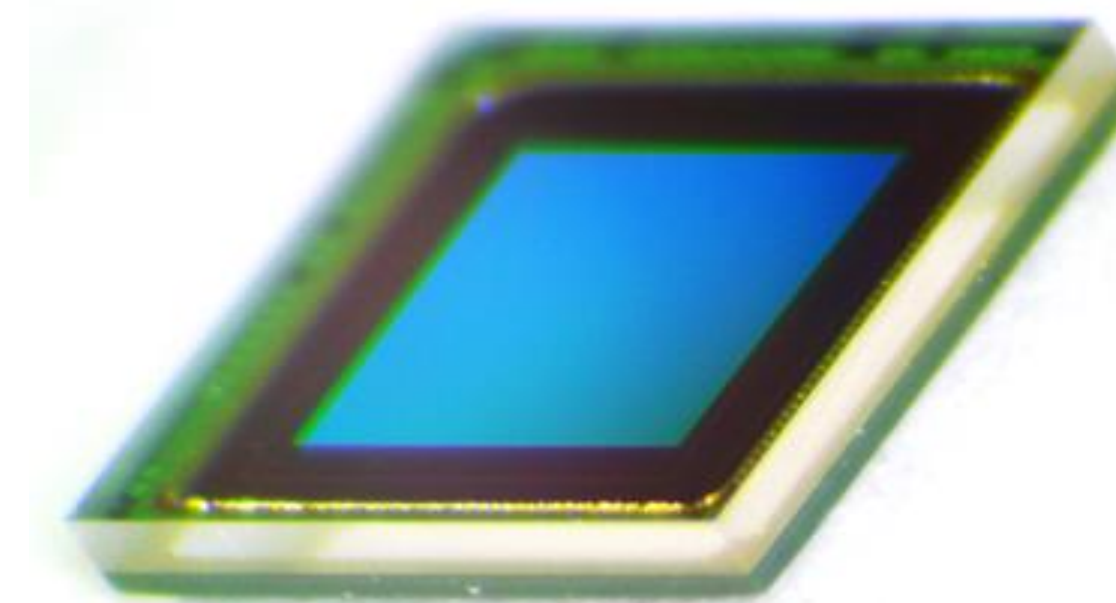


- Two silicon wafers bonded face-to-face, top CIS layer is thinned, light comes in from backside into photodiode (BSI)
- In-pixel metal-to-metal connection allows partition of pixel circuit on the two layers vertically
  - Front-end sensing circuits on CIS layer
  - ADC and memory on the bottom ADC layer
  - Small pixel size enabled by “hiding” ADC beneath photodiode
- Lower supply voltage on ADC layer to reduce power

# Implementation

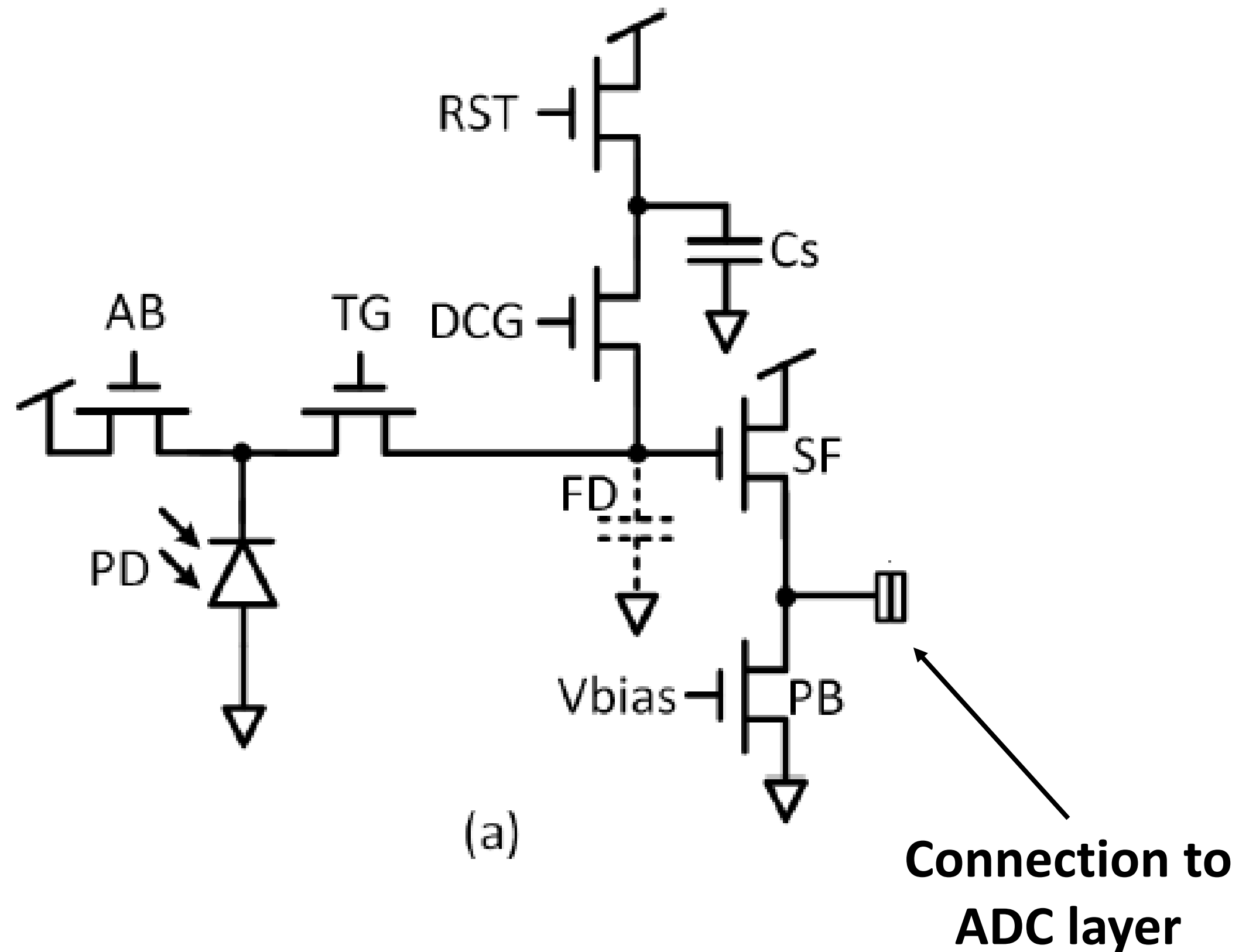


- 45nm(CIS layer)/65nm(ADC layer) stacked process
  - CIS layer is optimized for imaging performance
  - ADC layer is optimized for low power mixed mode circuit
- All ADCs operate simultaneously, inherently global shutter
- Specs:
  - Pixel size: 4.6um
  - Die size: 4 x 4 mm
  - Effective pixel: 512 x 512
  - Total power: ~5mW@30fps
  - Max frame rate: 480fps



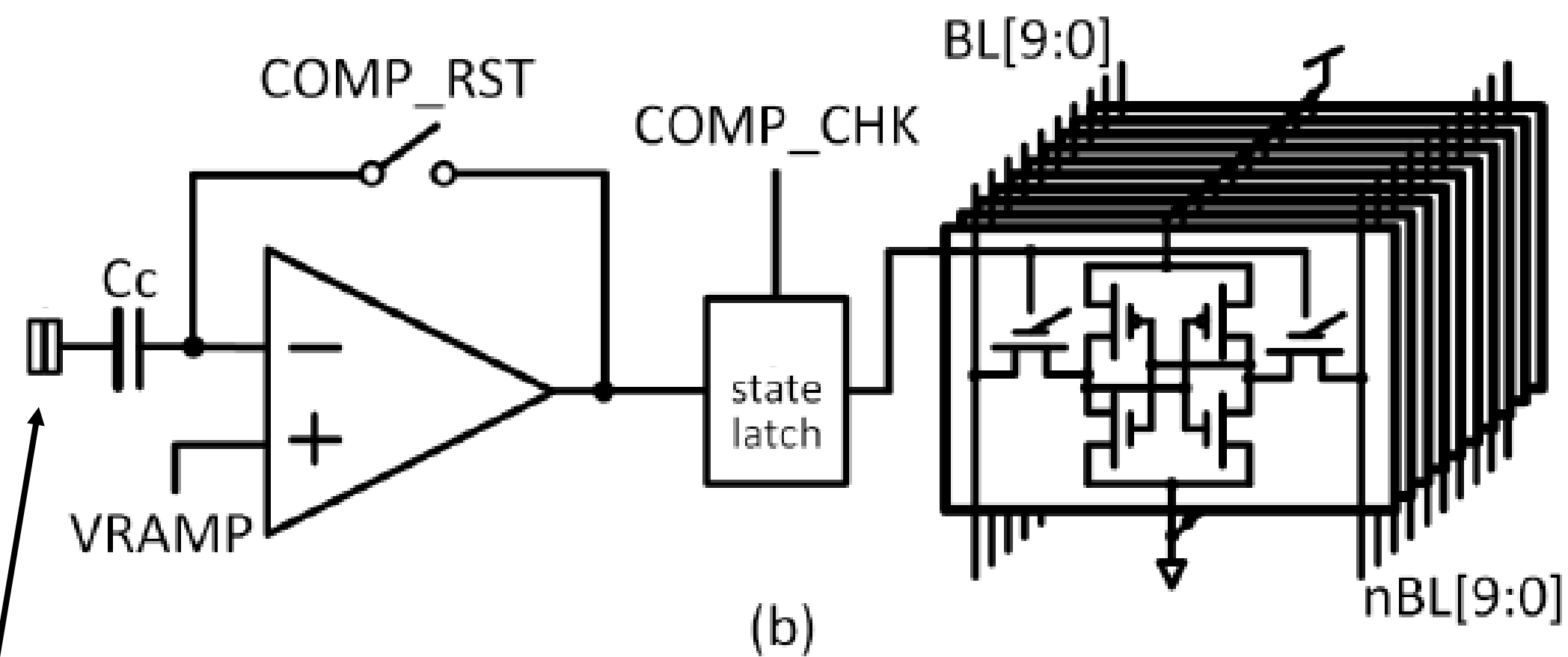
Chiao Liu, et. al., "A 4.6um, 512x512, Ultra-Low Power Stacked Digital Pixel Sensor with Triple Quantization and 127dB Dynamic Range", IEDM, December 2020

# CIS layer Circuits



- Global shutter and exposure controlled by AB gate
- Dual conversion gain enabled by DCG and Cs
  - High gain conversion when using FD cap only --- lower noise, but easier to saturate
  - Low gain conversion when using FD and Cs combined --- higher noise, higher dynamic range
- SF output connects to ADC layer

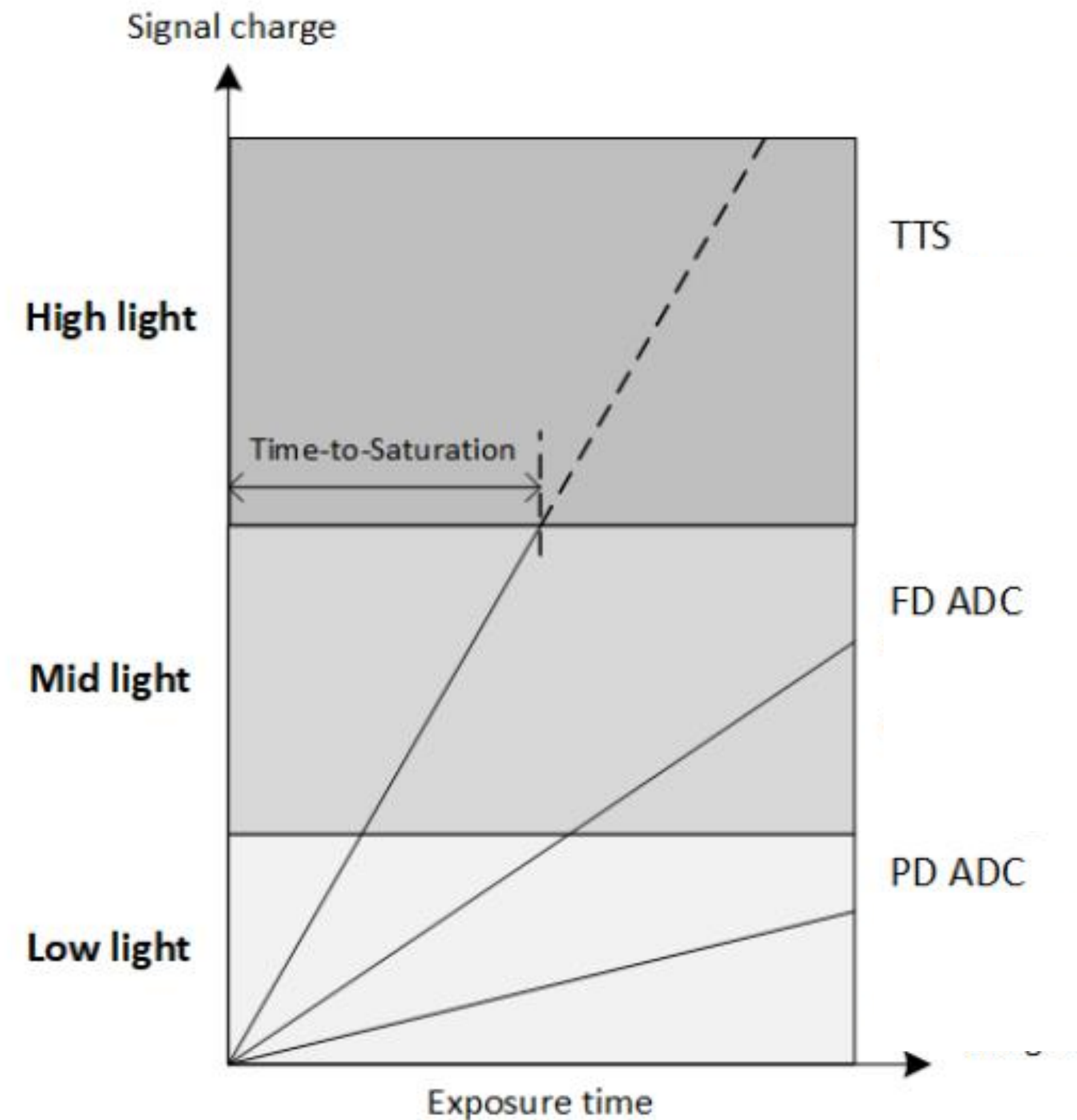
# ADC Layer Circuits



Connection from  
CIS layer

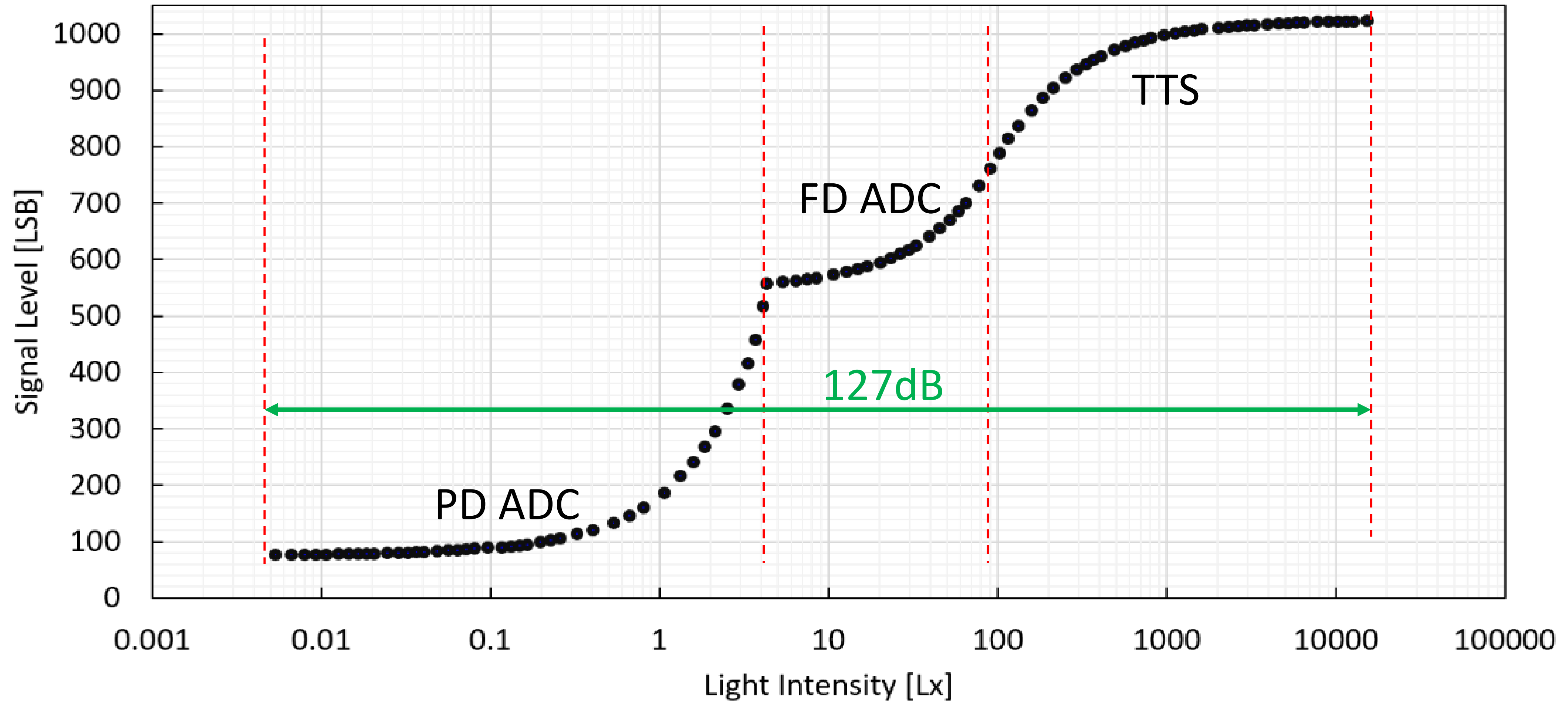
- Single-slope ADC
- CDS enabled by  $C_c$  and  $COMP\_RST$
- State latch controls the write state of pixel memory
- 10-bit 6-T SRAM per pixel
- ADC circuit in subthreshold, very low power
- Very short ( $\sim 20\mu s$ ) ADC time

# Triple Quantization Scheme

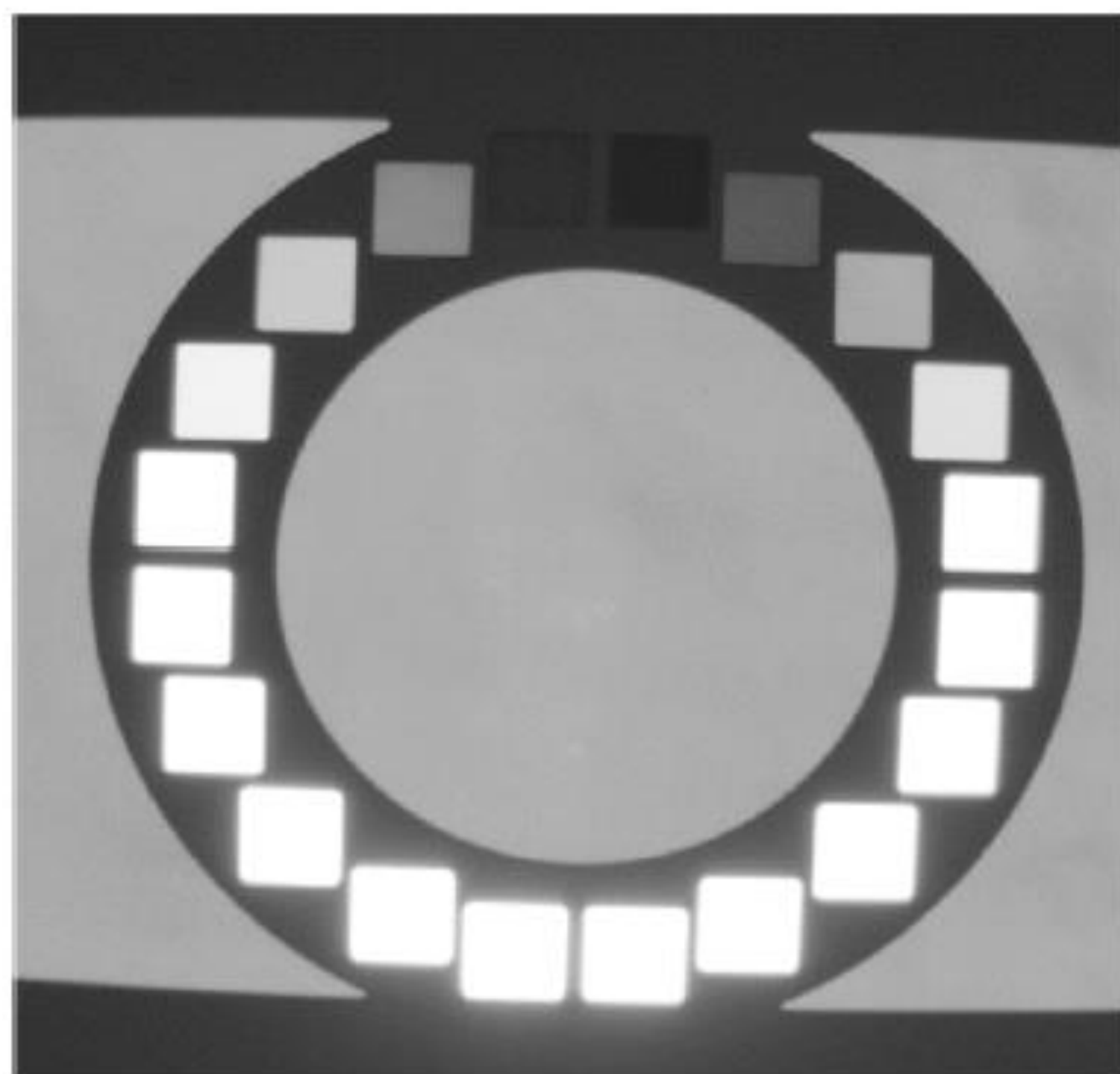


- Three ADC modes
  - Low light pixel --- linear high gain PD ADC
  - Mid light pixel --- linear low gain FD ADC
  - High light pixel --- time-to-saturation TTS mode
- Sequential operation within one exposure
  - TTS during exposure
  - PD ADC
  - FD ADC
- Each pixel “automatically” selects its optimal mode for its own light level

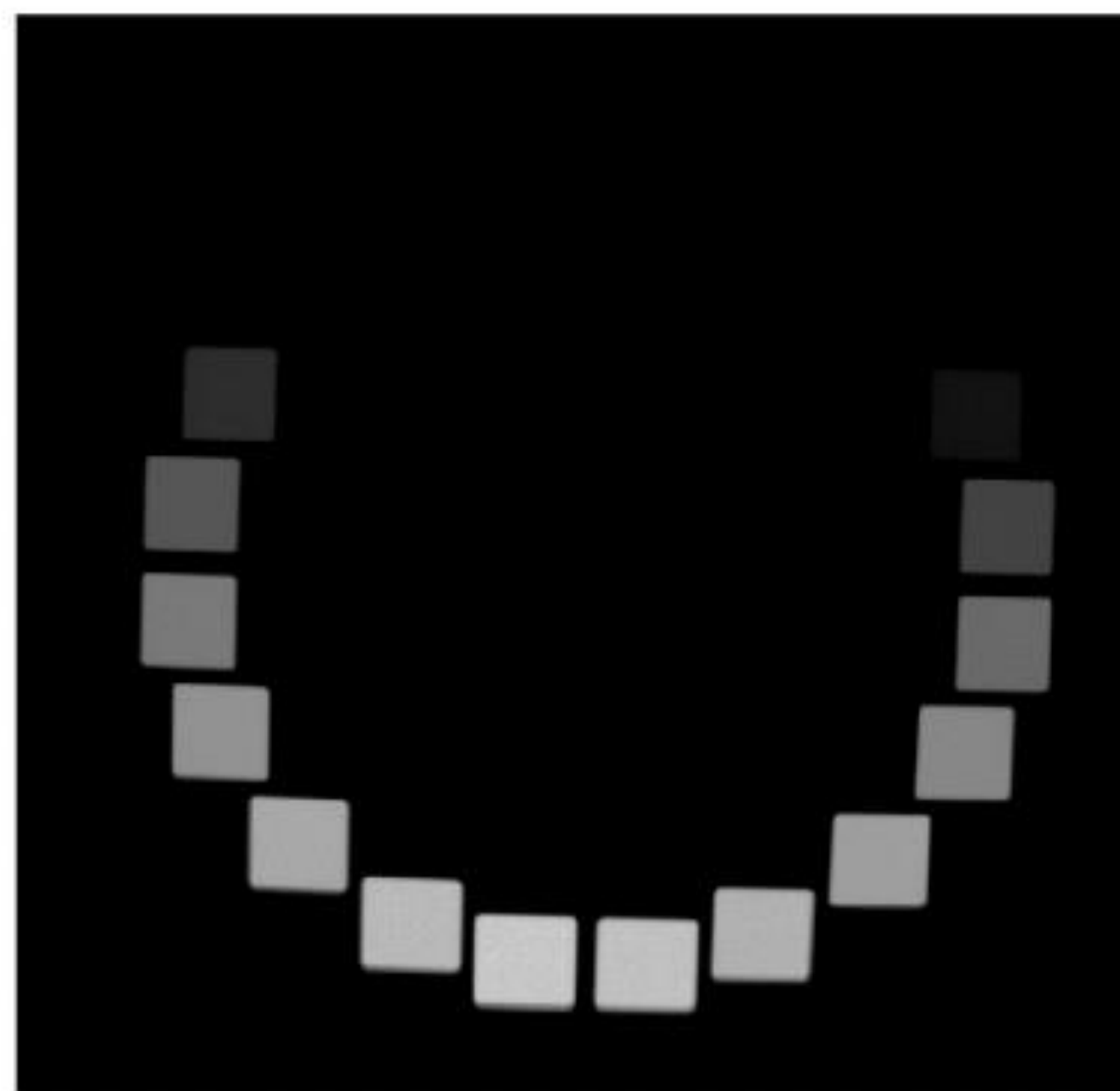
# Mapping Function from Light level to Digital Count with Triple Quantization



# HDR test Chart Image



(a)



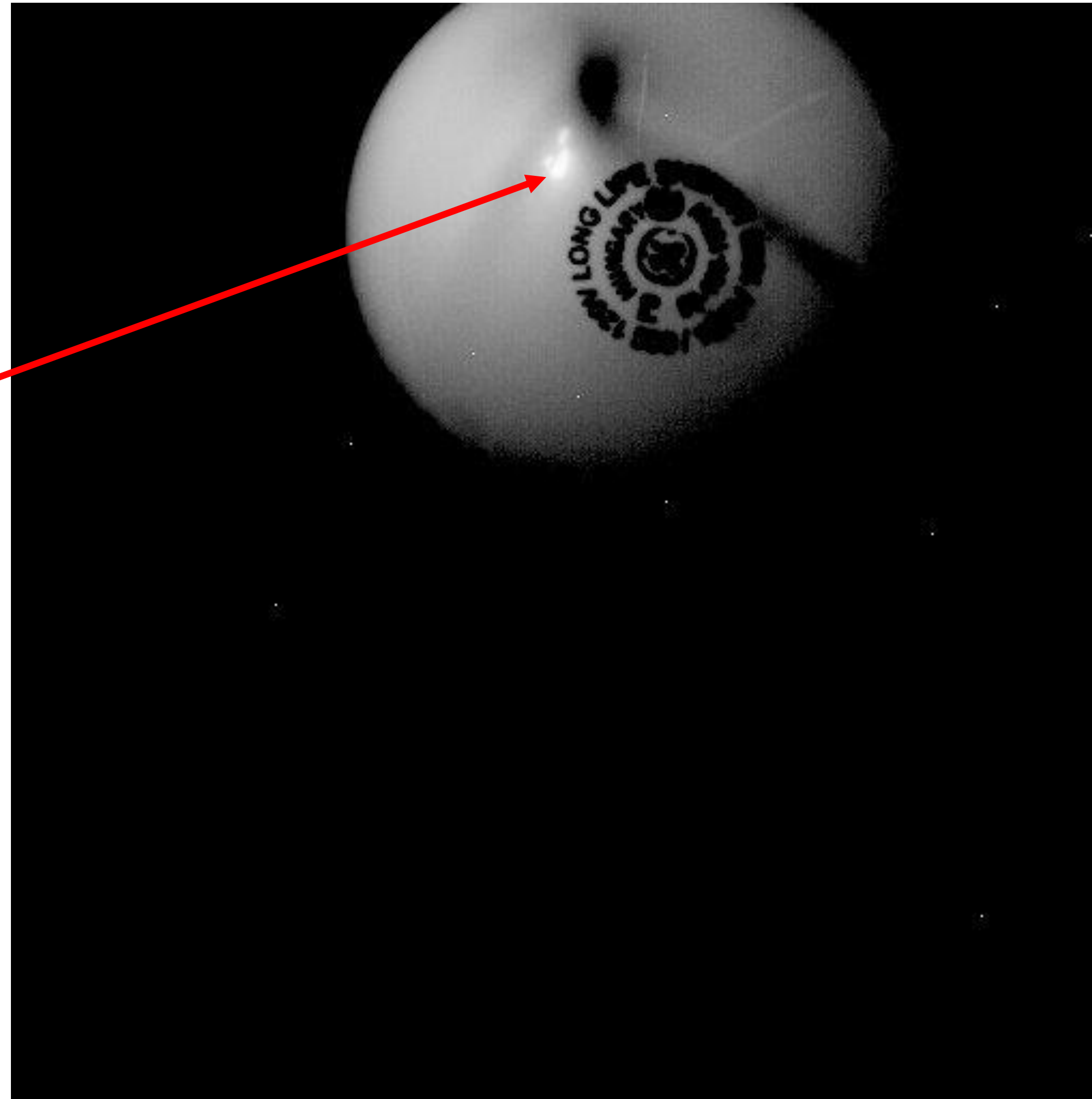
(b)

- Transmissive HDR test chart
  - 20 patches
  - Together cover 120dB DR
- Single exposure, single readout
- Two display setting for HDR rendering
  - (a) image clipped above 600
  - (b) image clipped below 500
- Gray scale different among patches demonstrated true 120dB capability



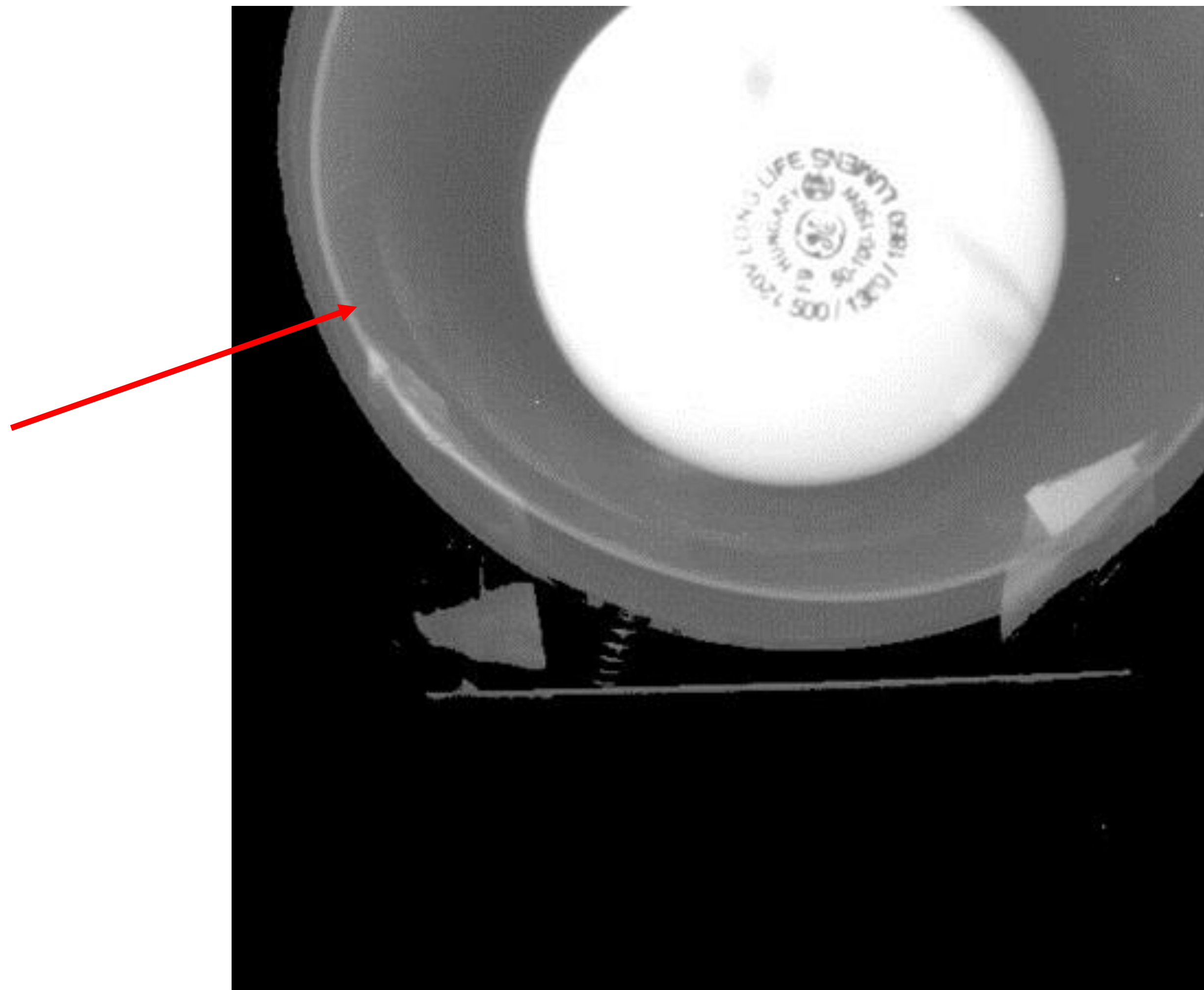
# HDR Image --- TTS Pixels

Filament



# HDR Image --- FD ADC Pixels

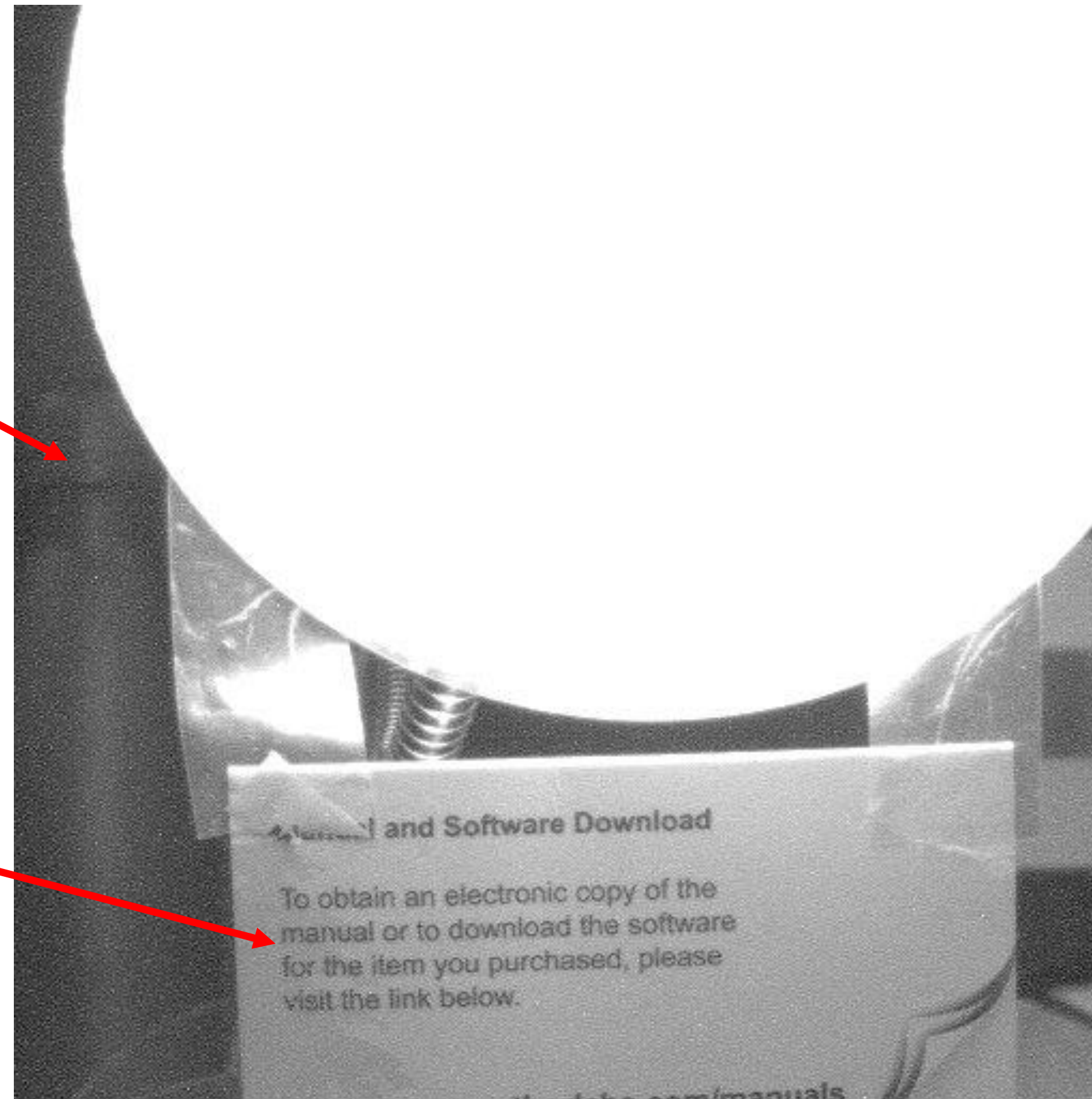
Lamp shield



# HDR Image --- PD ADC Pixels

Water bottle

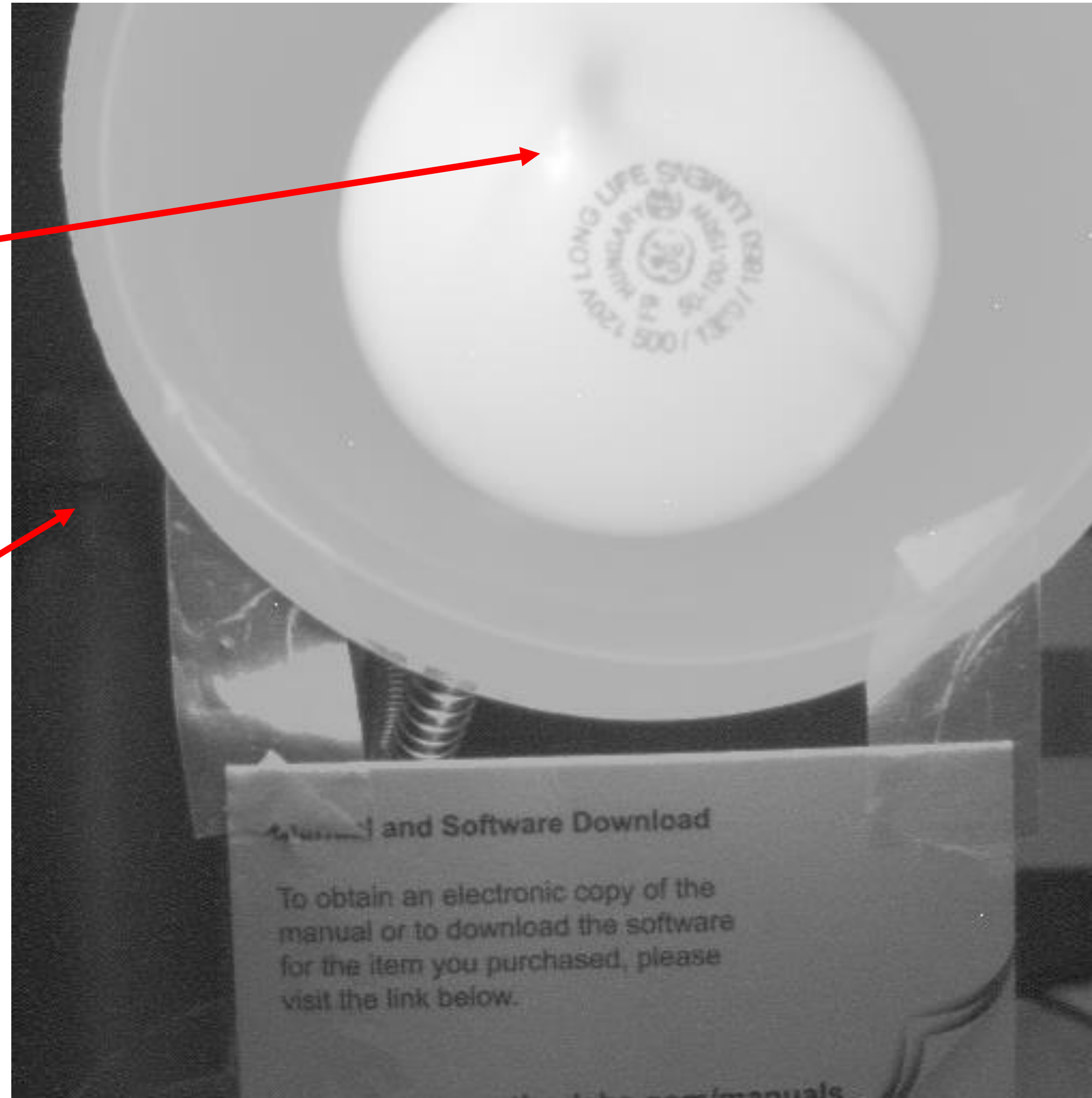
Letters on card



# HDR Image --- All Pixels

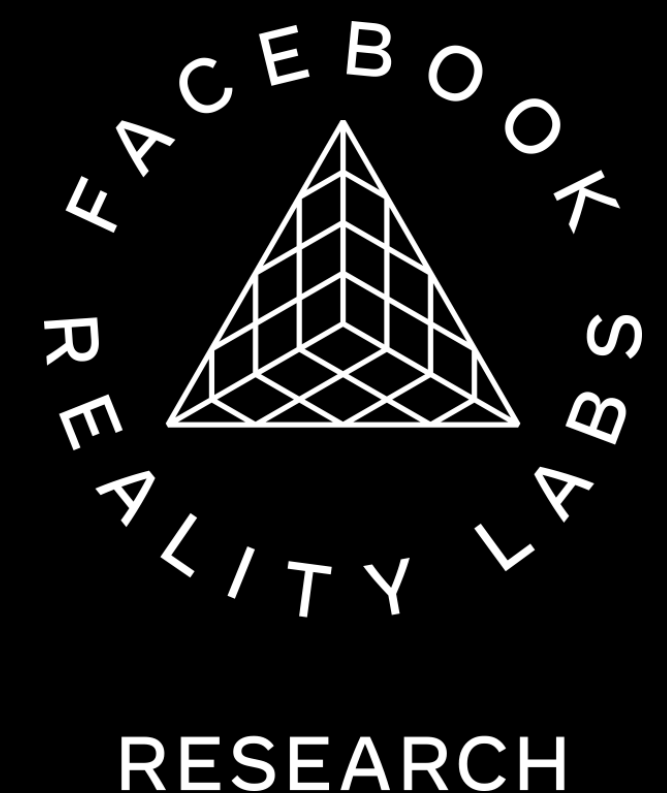
Filament

Water bottle

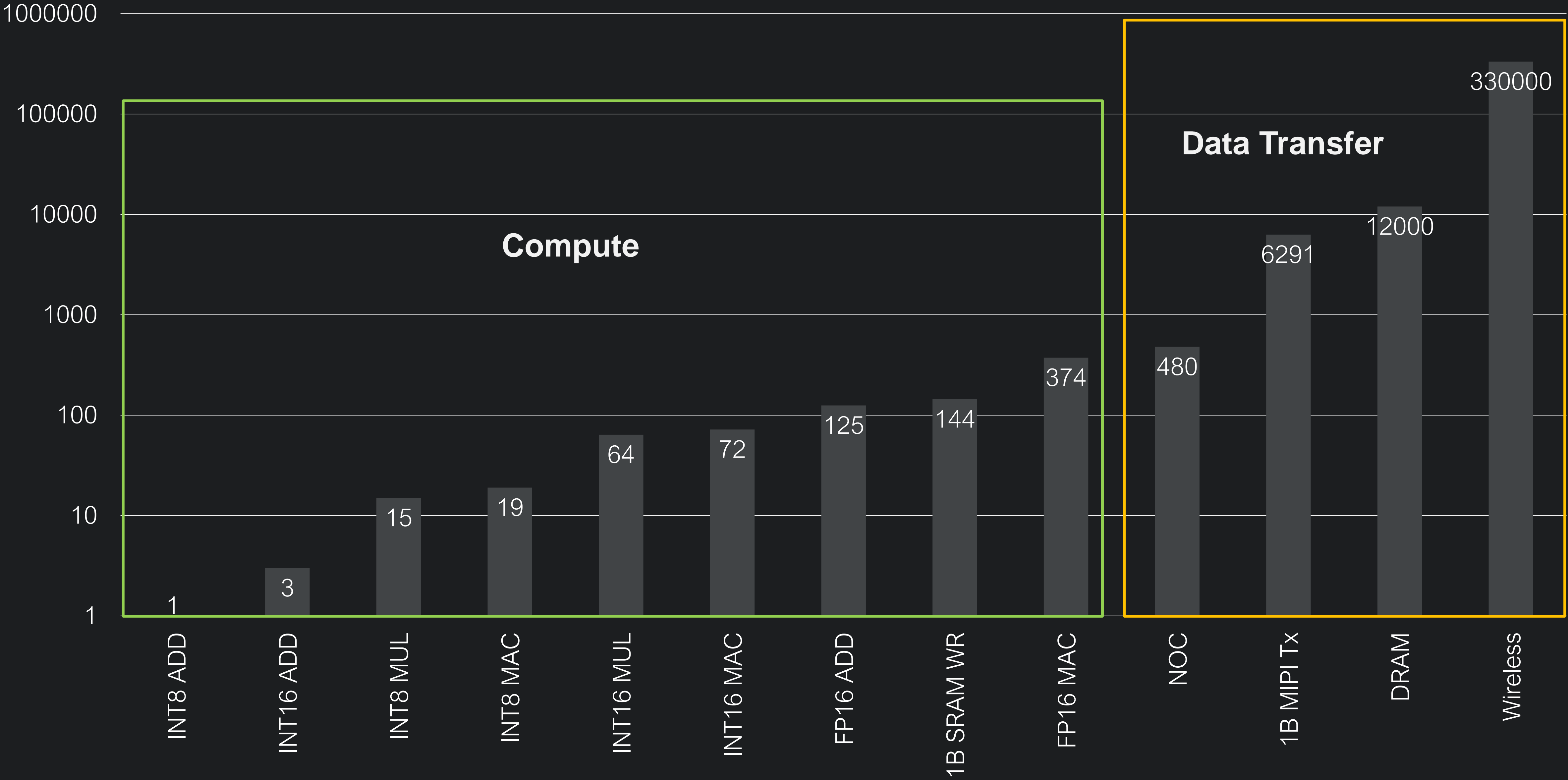


# Agenda

1. Image Sensing in AR/VR
2. CMOS Image Sensor Basics
3. A Digital Pixel Sensor w/ Ultra Wide Dynamic Range
4. A Distributed Sensing and Compute Architecture (IEDM 2019)

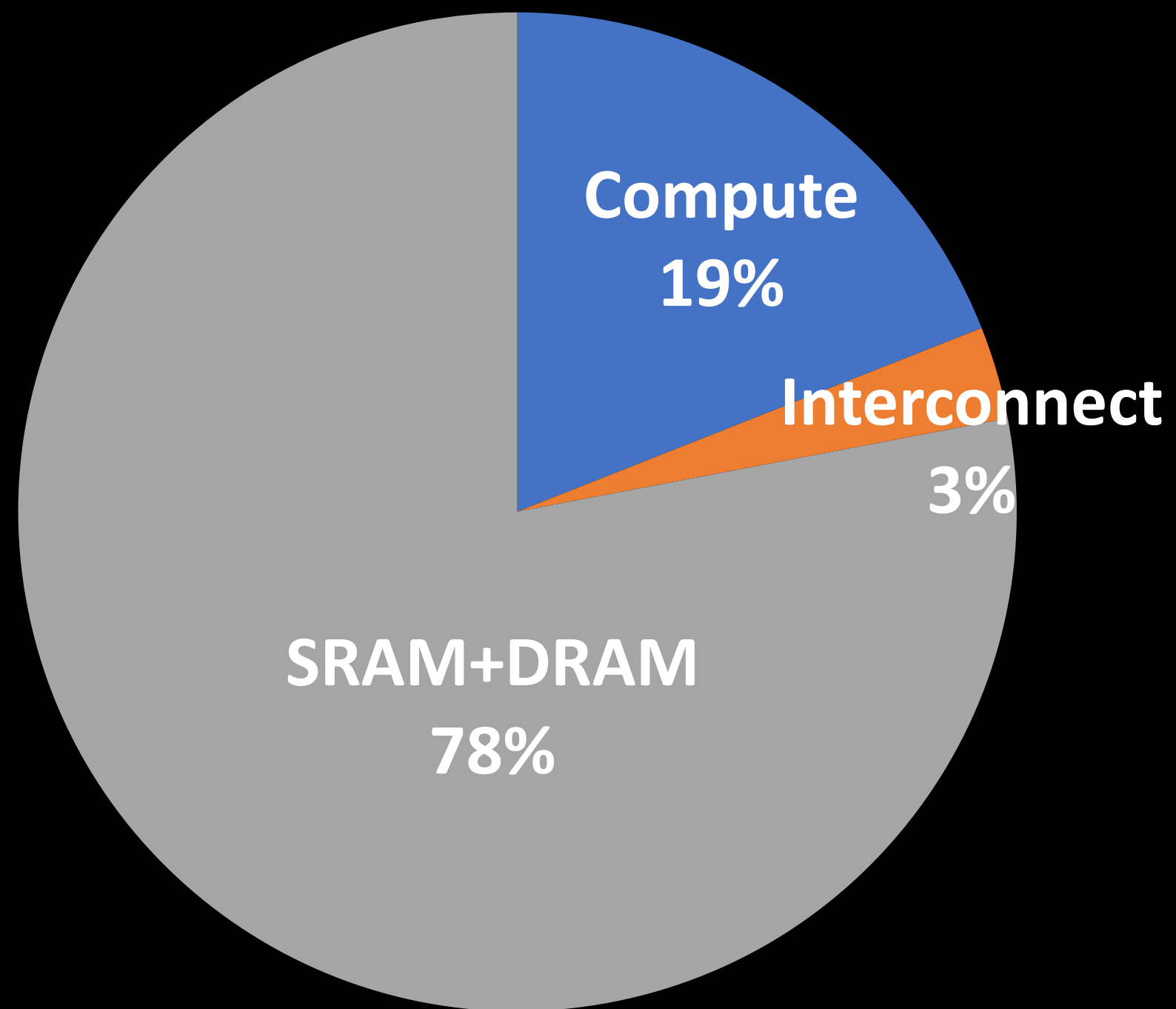


# Normalized Energy

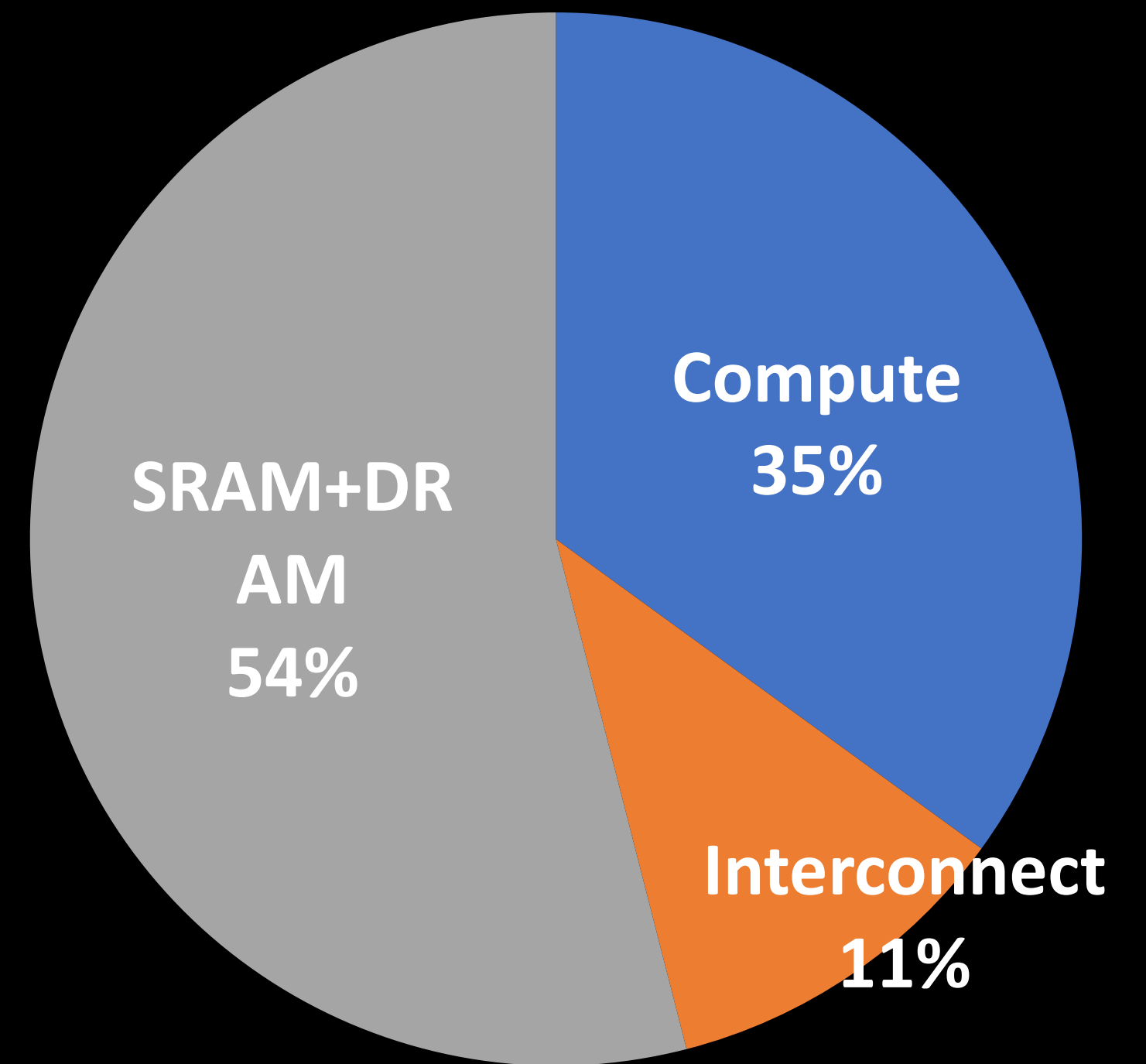


# Power Breakdown

---



SLAM



Hand Tracking

# System Level Optimization

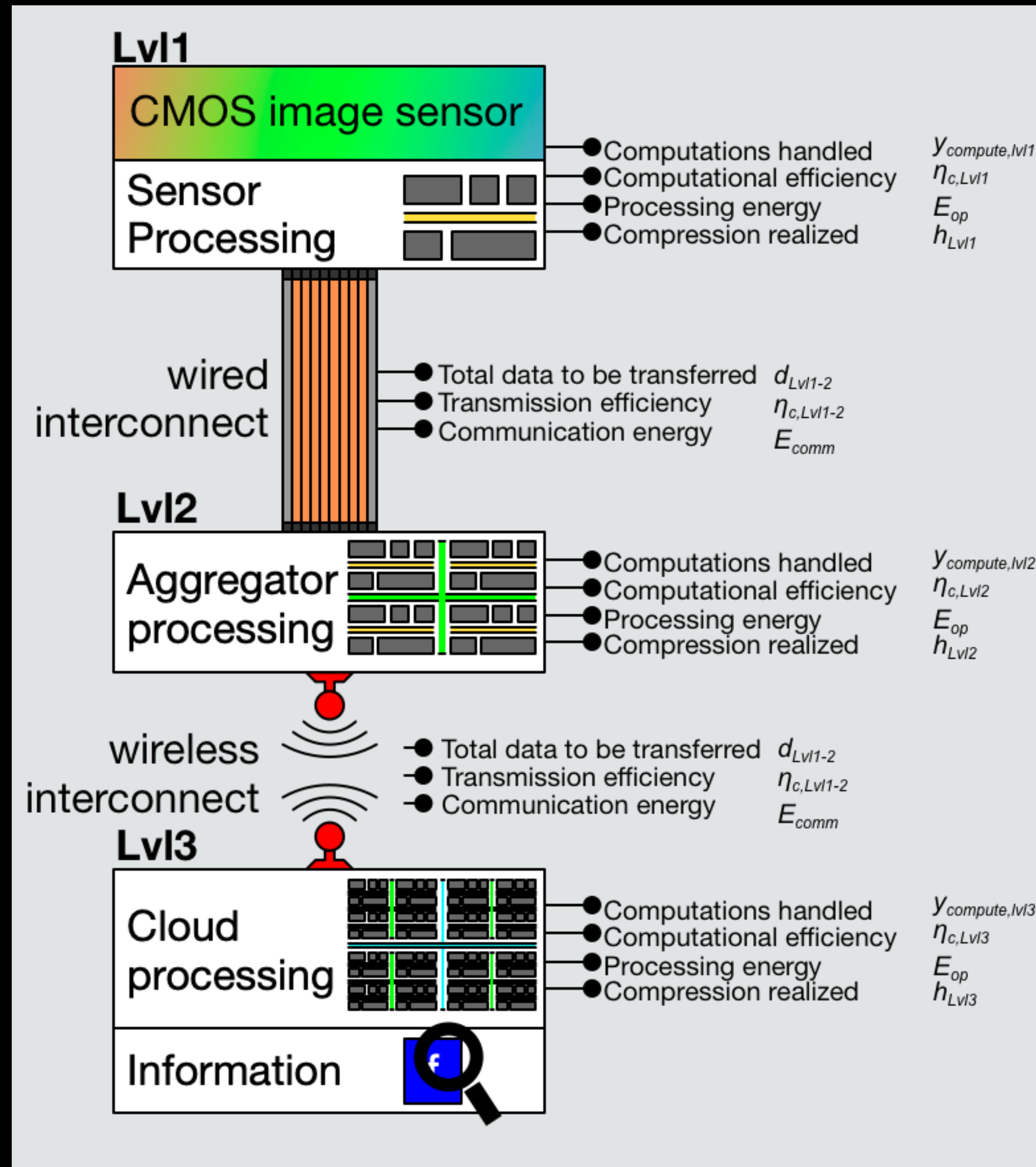
- Multi-camera configuration
- Overall latency requirement
- Communication cost
- Compute cost
- Algorithm optimization

→ Distributed sensing and compute architecture

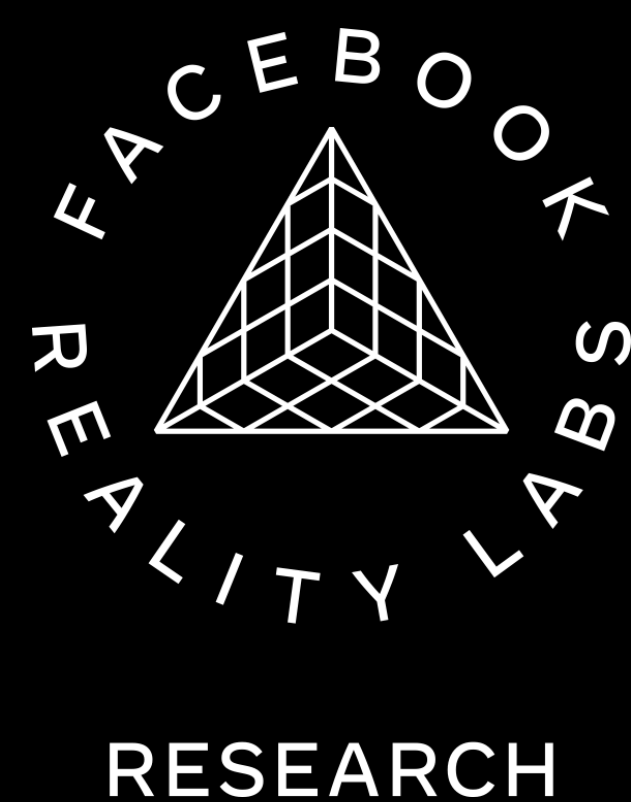




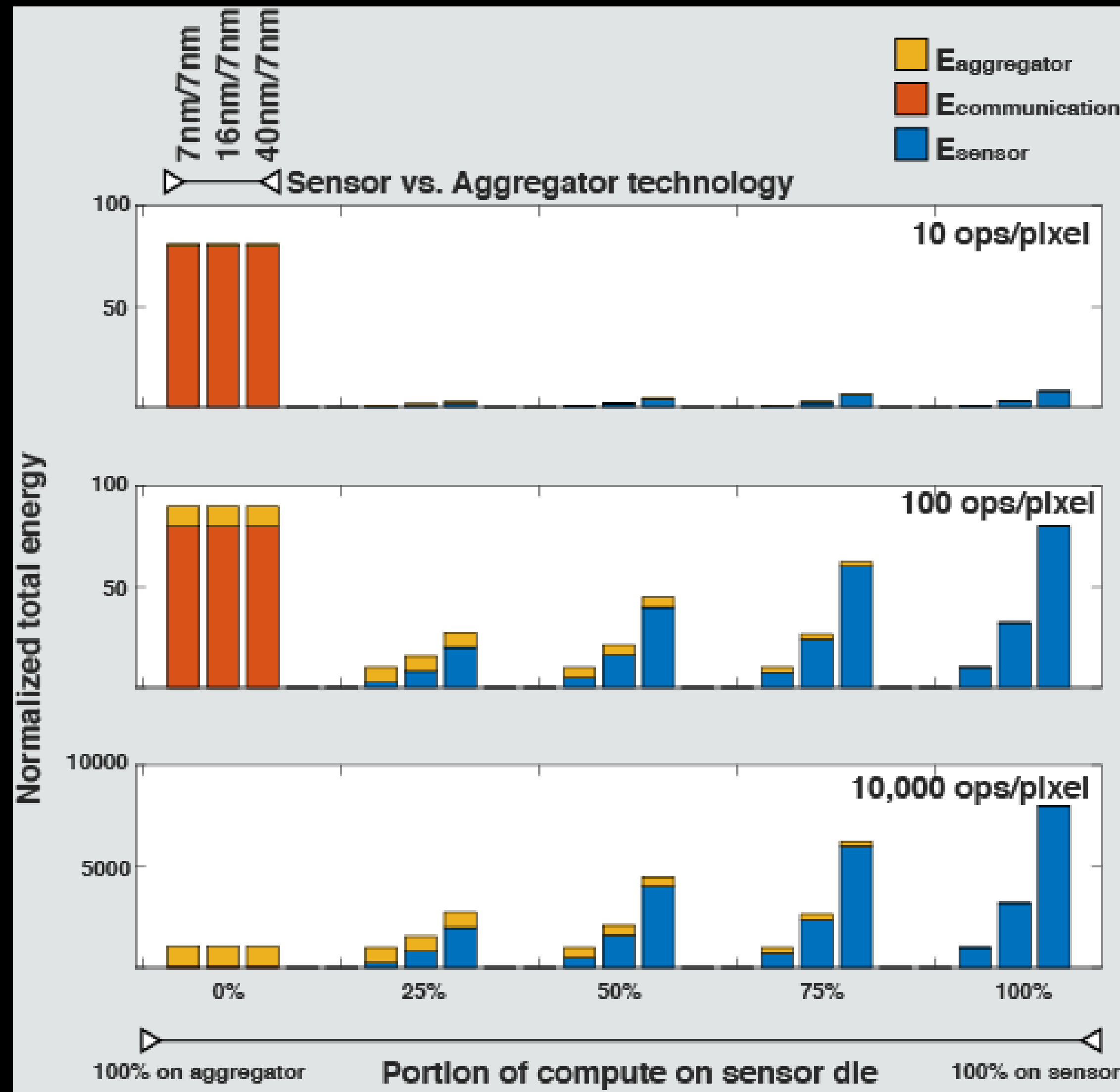
# A Distributed Sensing and Compute Architecture



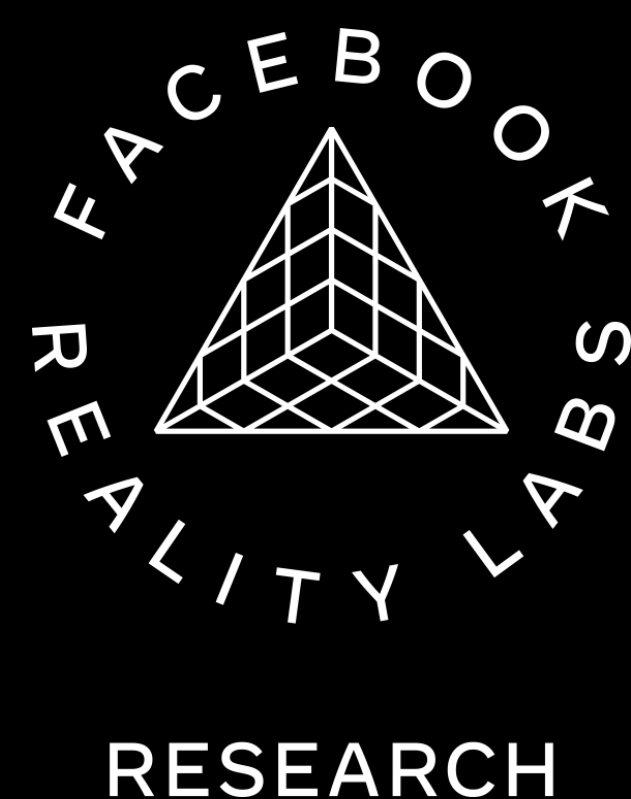
Chiao Liu, et. al., "Intelligent Vision Systems – Bringing Human-Machine Interface to AR/VR", IEDM, December 2019



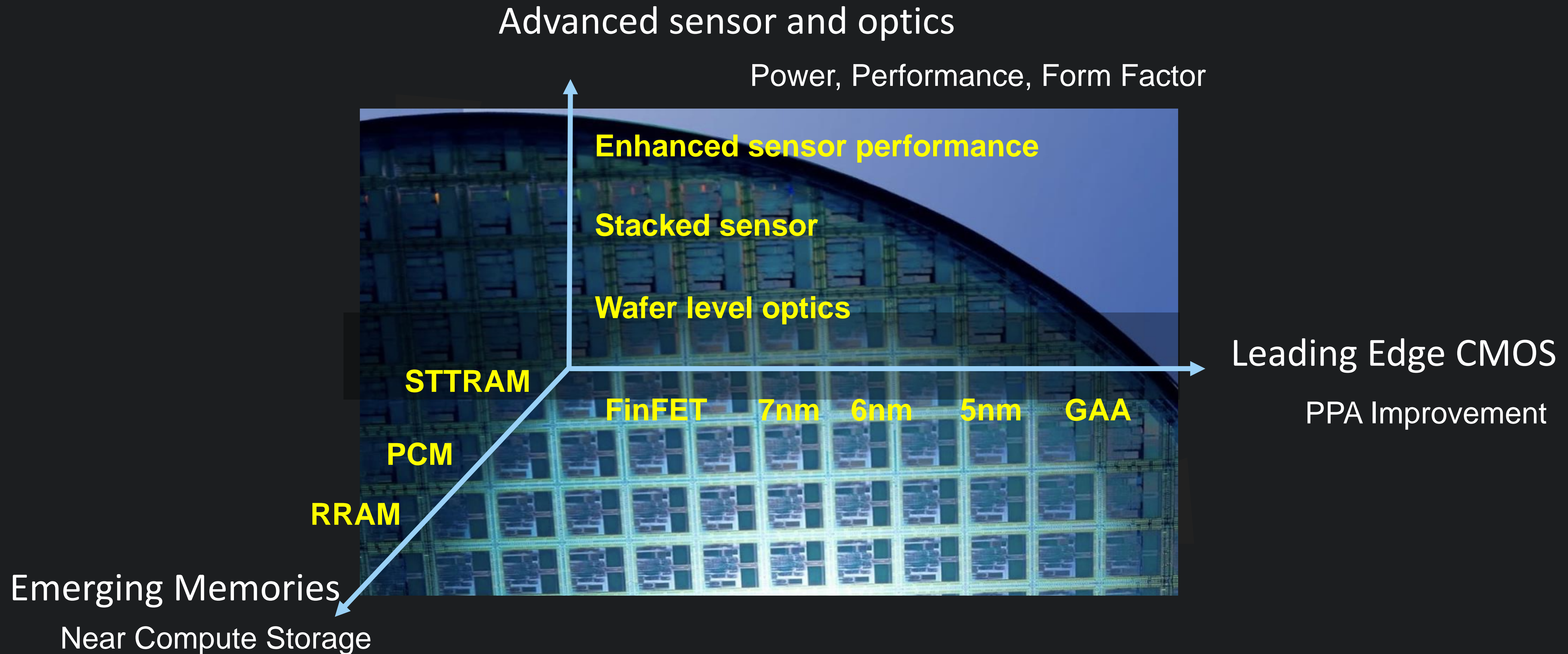
# Energies for different compute scenarios in a heterogeneous systems



Chiao Liu, et. al., "Intelligent Vision Systems – Bringing Human-Machine Interface to AR/VR", IEDM, December 2019



# Sensor and Silicon Technologies for AR/VR





NEW WAVE OF SENSING AND COMPUTING

Thank you

