CMOS Image Sensors: A Guide to Building the Eyes of a Vision System

Jon Stern, Ph.D.
Director, Optical Systems

Sept 2020

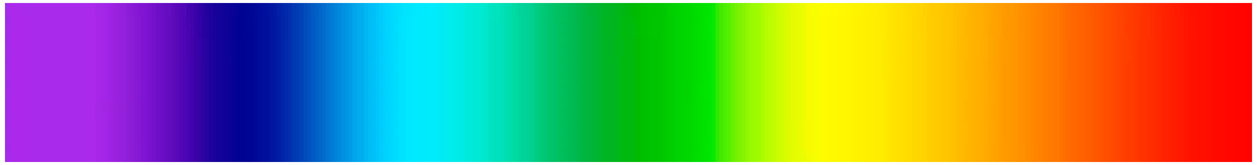
2020 embedded VISION summit

GoPro

Light Basics

What is Light?

- Light is represented as both a particle and an electromagnetic wave
  - When discussing sensors we tend to rather casually switch between the two
- A photon = a light particle
- The energy determines wavelength (higher energy = shorter wavelength)
  - Energy (J) = Planck’s constant (Js) x speed of light (ms^{-1}) / wavelength (m)
- When visible (to human beings), we perceive wavelength as color
- Intensity of light = number of photons



Bands of Light

Visible: 380-740nm

- Sensors: CCDs and CMOS image sensors (CIS)

Near Infra Red: 0.74µm – 1µm

- Sensors: CCDs and NIR-enhanced CIS

SWIR: 1-2.5µm

- Sensors: InGaAs sensors

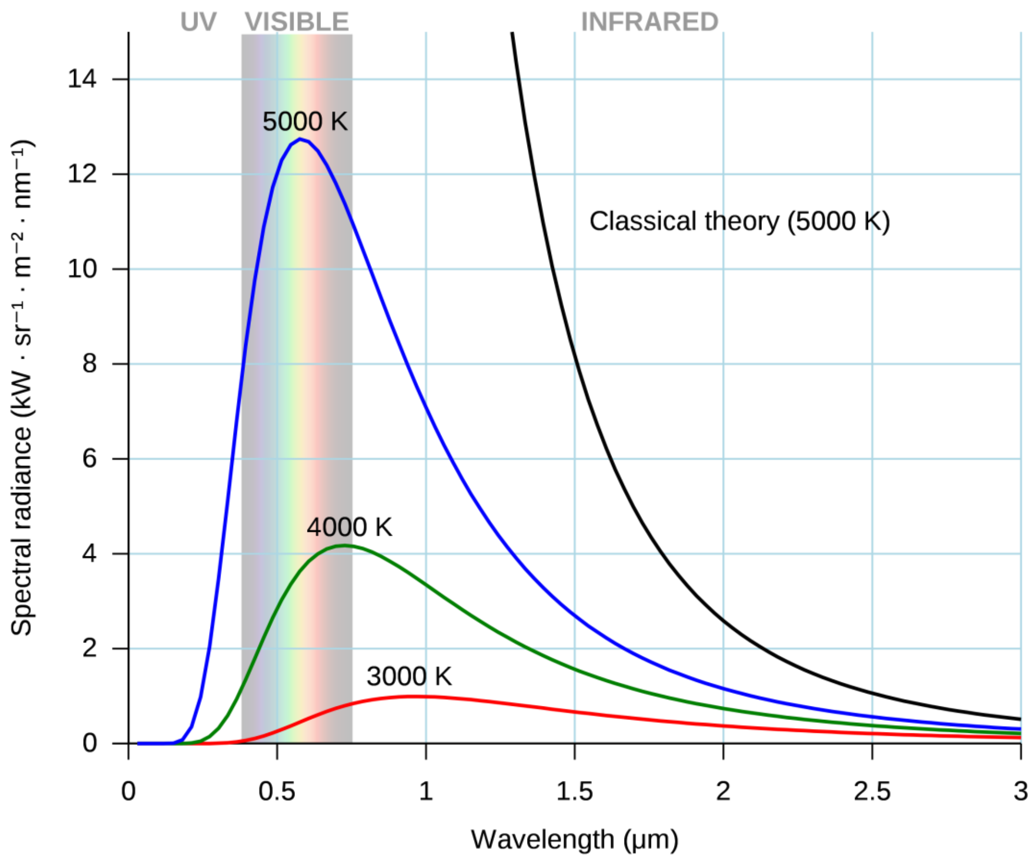
MWIR: 3-5µm

- Sensors: Indium Antimonide, Mercury Cadmium Telluride (HgCdTe), III-V semiconductor superlattices
- Thermal imaging

LWIR: 8-14µm

- Sensors: Microbolometers, HgCdTe
- Thermal imaging

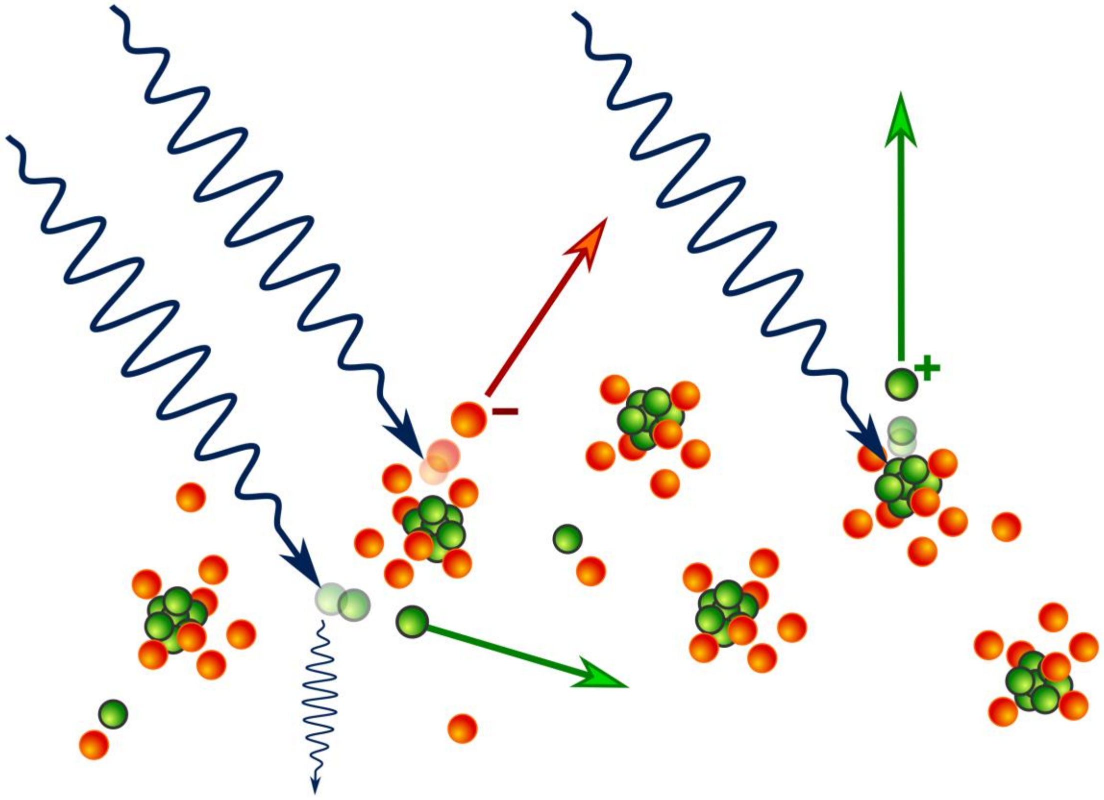
Black body radiation as governed by Planck's equation



By Darth Kule - Own work, Public Domain, https://commons.wikimedia.org/w/index.php?curid=10555337

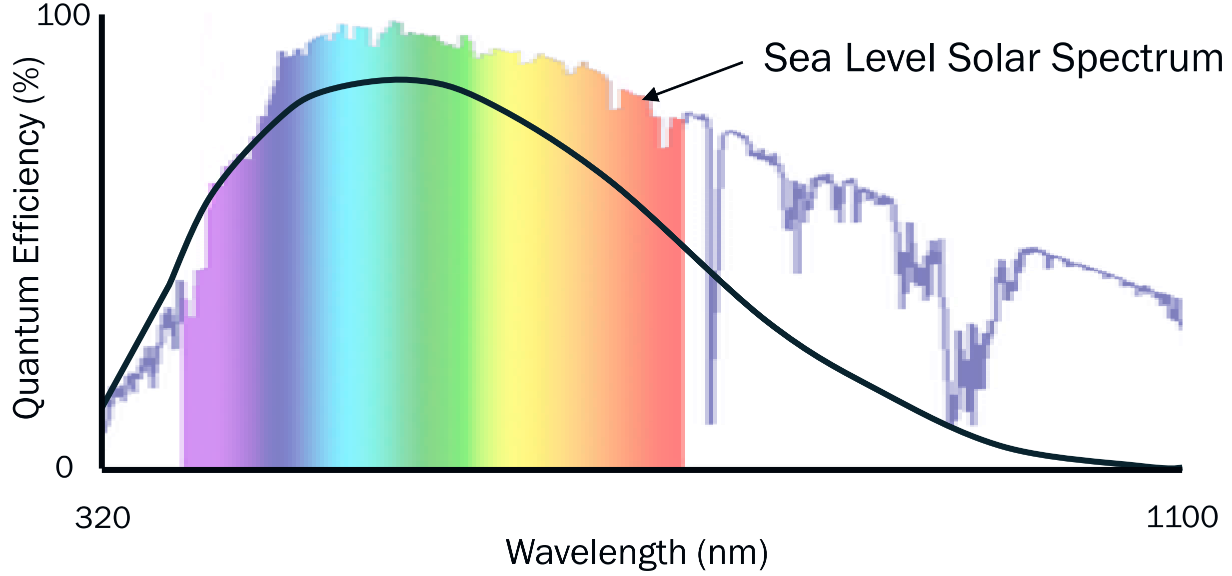
Photoelectric Effect

- Image sensors make use of The Photoelectric Effect
- Electromagnetic radiation hits a material and dislodges electrons
- The electrons can be collected and “counted”
- Interesting aside: Einstein was awarded a Nobel Prize for his work on the Photoelectric Effect, not for his theory of relativity
- The number of electrons is dependent on both light intensity and wavelength



Quantum Efficiency

- Quantum Efficiency (QE) is the ratio of number of electrons collected to the number of photons incident
- QE is sensor specific, and is frequently normalized (“Relative QE”) to obscure IP
- The peak QE of silicon aligns nicely with the peak response of human vision @555nm

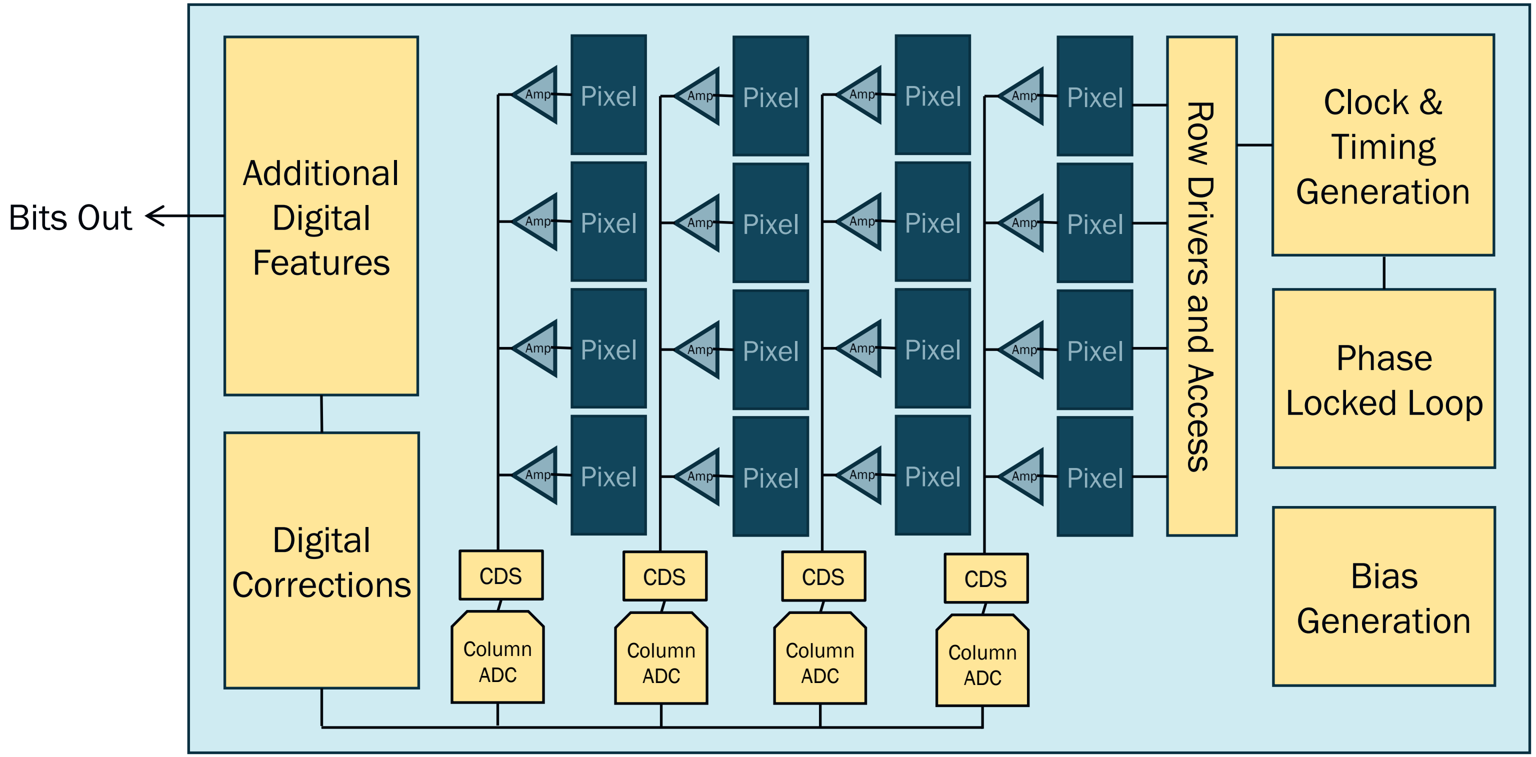


Typical Silicon Sensor QE

CMOS Image Sensors

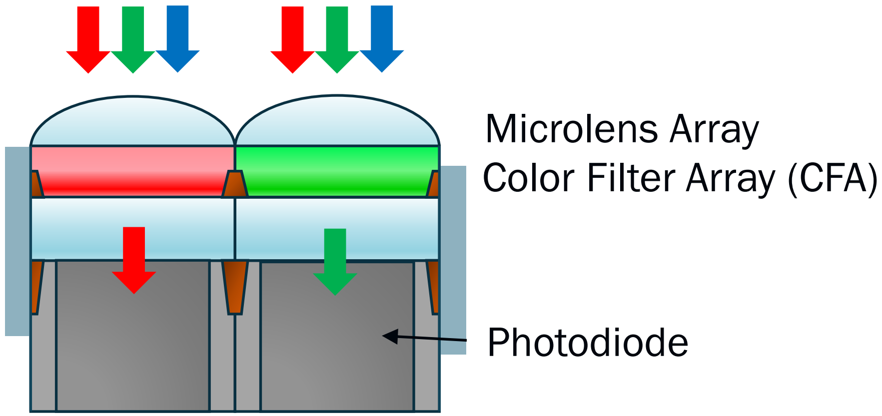
CMOS Image Sensor (CIS)

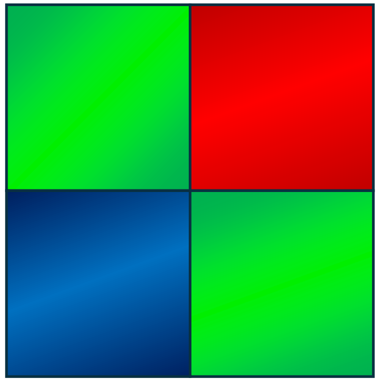
- Leverages Complimentary Metal Oxide Semiconductor (CMOS) manufacturing processes
- Each pixel contains a photodiode and a current amplifier (“active pixel sensor”)
- Very high level of integration (SOC)
  - On-chip ADCs, timing control, voltage conversion, corrections (black level, defective pixels, etc.)



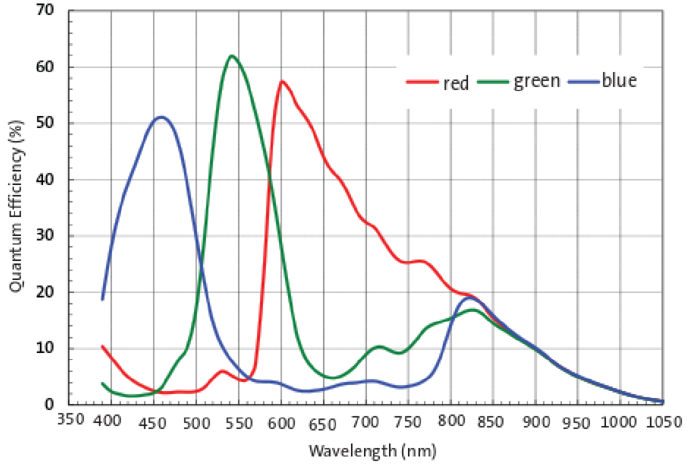
Seeing in Color

- Apart from in some specialist sensors (like Foveon) capturing color from a single image sensor requires spatial sub-sampling
- Place a color filter over each photodetector
- Demosaic algorithm interpolates missing colors to generate RGB for each location
- Add an IR cut off filter if you want to see “natural” (photoptic eye) color
- Novel filters patterns can be used for specialist applications, like CCCR (C= clear), RGB+IR (IR = infrared pass filter)





Bayer Color Filter Pattern



Quantum Efficiency for Color Sensor

Noise Sources in CMOS Image Sensors

Shot noise

- Noise that is intrinsic to the quantum nature of photons
- Uncertainty in measurement
- Shot noise = SQRT (number of photons)

Read noise

- RMS noise of sensor circuitry in electrons

Fixed pattern noise

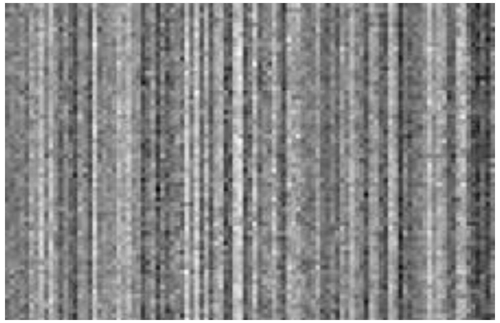
- Pixel FPN – pixel to pixel dark offsets
  - Includes dark signal non-uniformity, which changes with integration time and temperature
- Column FPN - per column dark offsets
  - Visible well below the pixel temporal noise floor (e.g., visible down to ~1/20th read noise)
- Row FPN – per row dark offsets

Photo Response Non-Uniformity

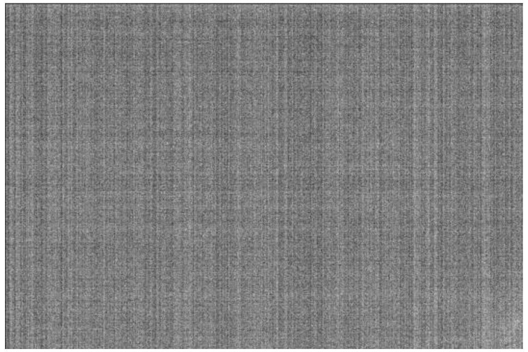
- Pixel to pixel variation in responsivity (gain error)

Row noise

- Per row variation in the black level of each row. Dominated by temporal noise (row FPN generally not an issue in modern CIS)



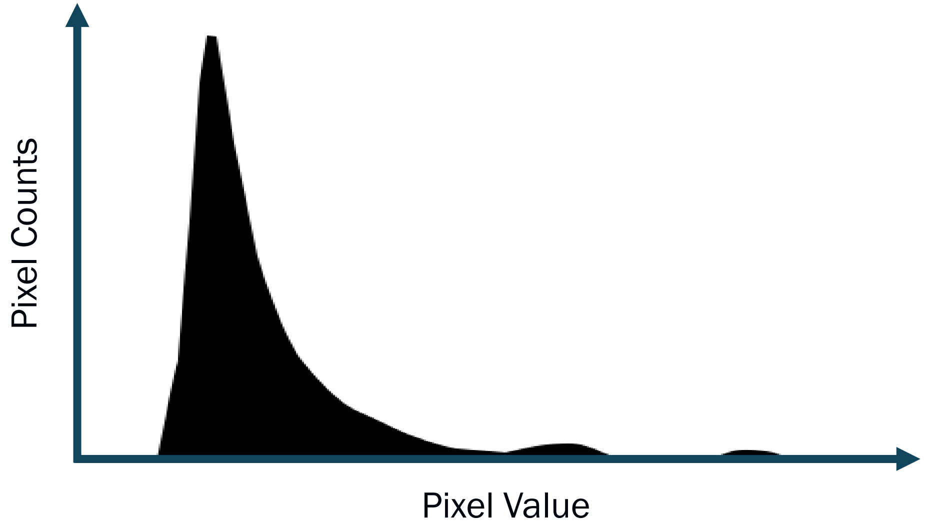
Column fixed pattern noise

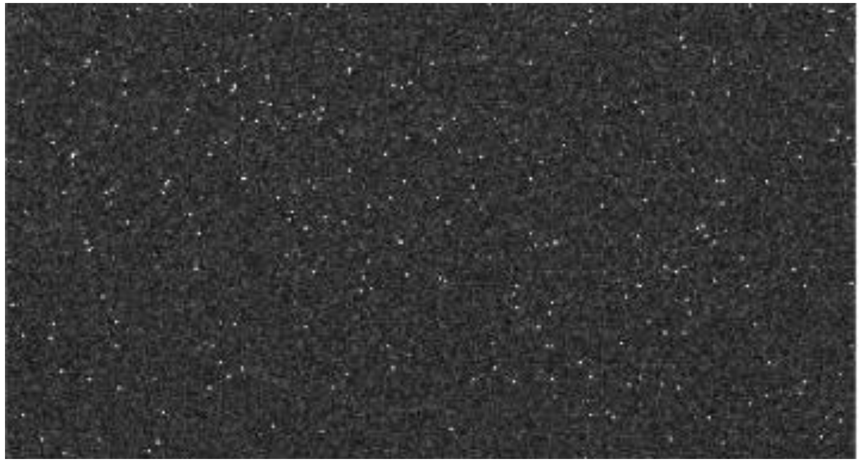


Column fixed pattern noise + row noise

Hot Pixels

- Pixels with high leakage
- Bright pixels in images captured in the dark with long exposure and/or high temperature
- Need correcting (really hiding) by the image processing pipeline
  - Static map programmed at factory
  - Dynamic detection and correction
  - Replace with average of neighboring pixel values (in the same color plane)



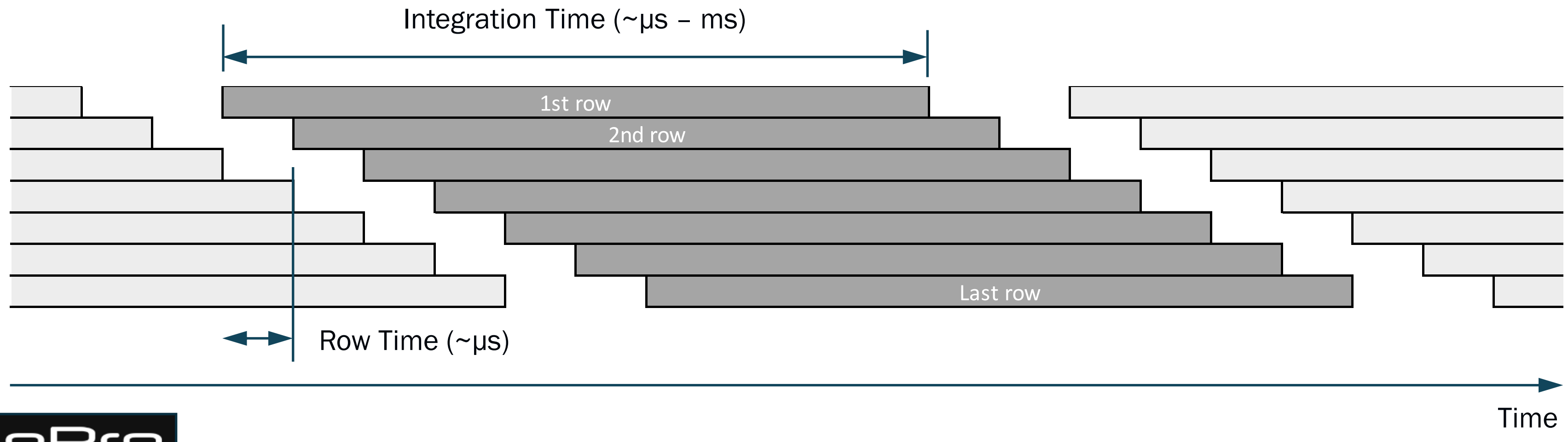


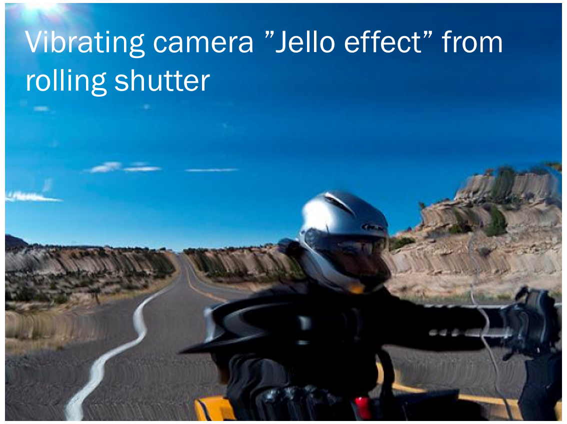
Global vs Rolling Shutter

Global shutter: all sensor rows capture light at the same time

Rolling shutter:

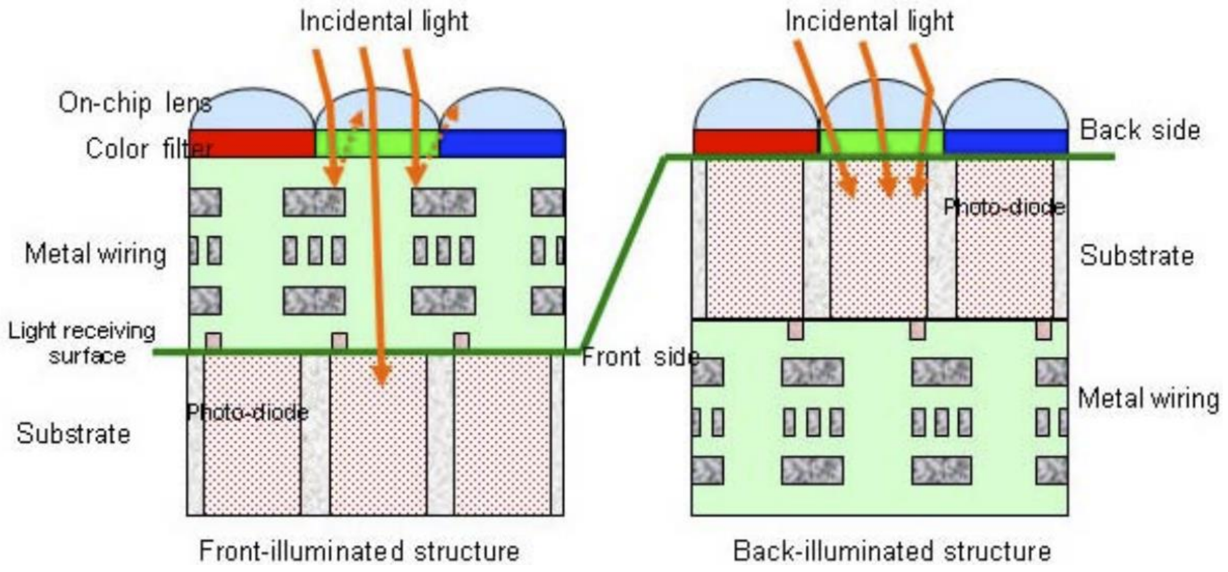
- Each row reset one at a time, then read out n row periods later
- Creates rolling shutter artifacts (moving object distortions, “jello effect” from camera vibration)





Front-Side Illumination vs Back-Side Illumination

- Front-side illumination (FSI) sensors have the photodiode below the metal wiring
  - Impacts sensitivity, acceptance angle, optical cross talk (photons being captured by the wrong photodiode)
- In back-side illumination (BSI) sensors the wafer is flipped and bonded to a blank wafer, or an ASIC layer, and thinned
  - The color filter array and microlenses are built on top of the thinned photodiode layer



In FSI (left) metal impedes the optical path to the photodiode. In BSI (right), the optical stack height is minimized

Key Sensor Specifications Impacting Image Quality

- Full well capacity (really linear full-well capacity)
  - Small pixels: 5,000 electrons
  - Medium pixels: 10,000 electrons
  - Large pixels: 50,000 electrons
- Sensitivity
  - Voltage swing per lux-s
- Read noise
  - The noise floor measured in electrons
  - Small-pixel, low noise sensor have read noise ~2 electrons
- Resolution
  - Resolution is really a measure of linear resolving power, but colloquially it has come to mean number of pixels
  - Higher resolution means higher spatial sampling of the optical modulation transfer function (MTF)

Image Quality Metrics

- Signal to noise ratio
  - Ratio of the signal from photons collected to unwanted noise
  - SNR10: the number of lux (measurement of luminous flux per unit area) to achieve an SNR of 10
- Dynamic range
  - Ratio between the max. output signal level and noise floor
  - Sensor companies typically misrepresent this as:
    - Max. output signal at minimum analog gain / noise floor at max. analog gain
    - This is unachievable within a single frame
  - $dB = 20 \times \log \left(\frac{Brightest\ Lux}{Darkest\ Lux}\right)$
- Resolution
  - Not just about number of pixels. More formally is a measure of the smallest features that can be resolved. Described in more detail later

Illuminance (Lux)	Scene
0.2 - 1	Full moon
100	Dim indoor
500 – 1,000	Office
10,000– 30,000	Outdoors, cloudy
100,000	Direct sunlight

Ratio of Brightest/Darkest	Dynamic Range (dB)
10	20
100	40
1,000	60
10,000	80
100,000	100

Image Sensor Interfaces

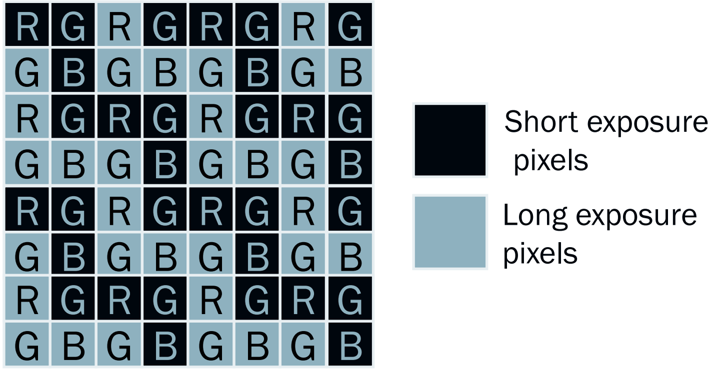
- MIPI
  - D-PHY
    - Dominant interface in phones for a decade
    - Up to 4 differential paired lanes
    - Separate clock lane
  - C-PHY
    - Higher performance interface with embedded clock
    - 3-wire lanes (aka “trios”)
    - Starting to gain traction in high performance phones
    - Likely to gain broader adoption over time
  - A-PHY
    - New MIPI standard for automotive adoption
- Various proprietary LVDS, subLVDS, and SLVS interfaces
- It is important to either select a sensor that will directly interface with your SOC, or accept the added cost and power of an external interface bridge chip

High Dynamic Range (HDR) Sensors

Typical commercial sensors have a dynamic range around 70dB

HDR sensors use a multitude of techniques to extend the dynamic range that can be captured

- Multi-frame HDR: capture sequential frames with long and short exposure times
  - Long exposure captures the shadows
  - Short exposure captures the highlights
  - Stitch the frames together to create a single higher dynamic range frame with higher bit depth
  - For scenes with motion, this technique can result in significant motion artifacts (ghosting)
- Overlapping multi-frame exposure
  - Multiple exposures are produced on a line-by-line basis, rather than frame by frame
  - Greatly reduces the motion artifacts by minimizing the time delta between the different exposure read outs
- Spatially interleaved methods
  - Multiple exposure times are captured simultaneously by varying the exposure time of different groups of pixels (alternate lines, zig-zag pattern, etc.)
  - Almost eliminates HDR motion artifacts, but reduces resolution and creates aliasing artifacts on high contrast edges



Sony SME HDR

High Dynamic Range (HDR) Sensors - Continued

- Dual conversion gains
  - Technology that allows the pixel gain (µV/e-) to be switched between two values
  - Can be used in a multi-frame, or spatially interleaved manner with similar characteristics, but improved noise (as there is no short exposure throwing away photons)
- Logarithmic pixels
  - Pixel has a logarithmic response, or a logarithmic response region, that greatly extends dynamic range compared to a linear mode pixel
  - Mismatches between the pixel response curves and changes with temperature can create high fixed pattern noise and high photo-response non-uniformity
- Overflow pixels
  - Pixels that contain an overflow gate that allow charge (electrons) to flow to a secondary overflow node
  - Requires additional per-pixel storage node, and control line mean that the technology only works for larger pixels (~6µm)

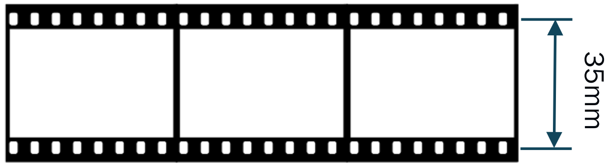
Optics for Image Sensors

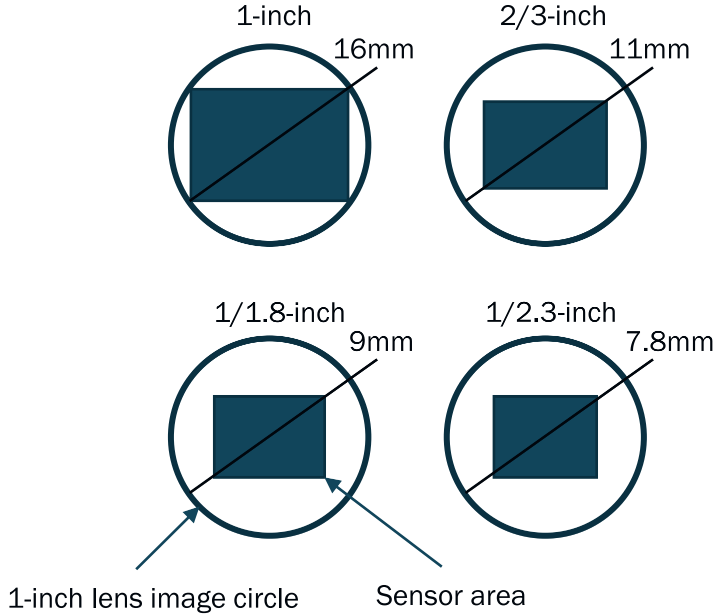
Matching Optics with a Sensor: Overview

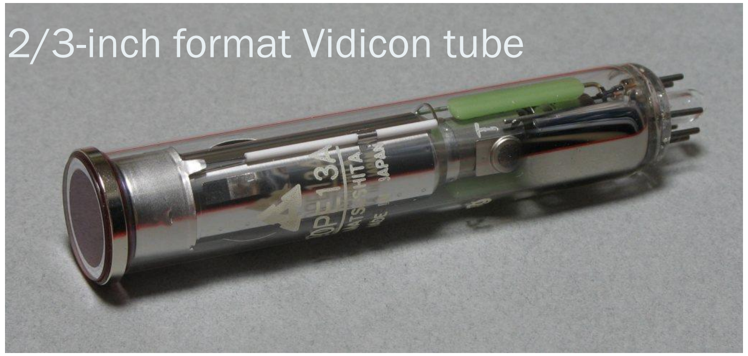
- Optical format
- Chief ray angle
- Focal length
- Magnification
- Field of view
- F-number
- Depth of field
- Wavelength range
- Modulation transfer function
- Relative illumination
- Stray light

Optical Format

- Optical format is probably not what you think!
- A 1/2.3-inch format sensor does not have a diagonal of 1/2.3” (11.04mm). It’s actually 1/3.2” (7.83mm)
- Convention is legacy of imaging tube technology, where the stated format was the mechanical diameter of the tube
- Optical format ~3/2 x sensor diagonal
- From film cameras, “35mm format” is 36mm x 24mm (43mm diag.)
    - 35mm is the pitch between the feed sprocket holes
- The lens and sensor optical formats must be matched
- Smaller sensor optical format is okay, but this will crop the field of view, and may create a CRA mismatch issue (more on this later)

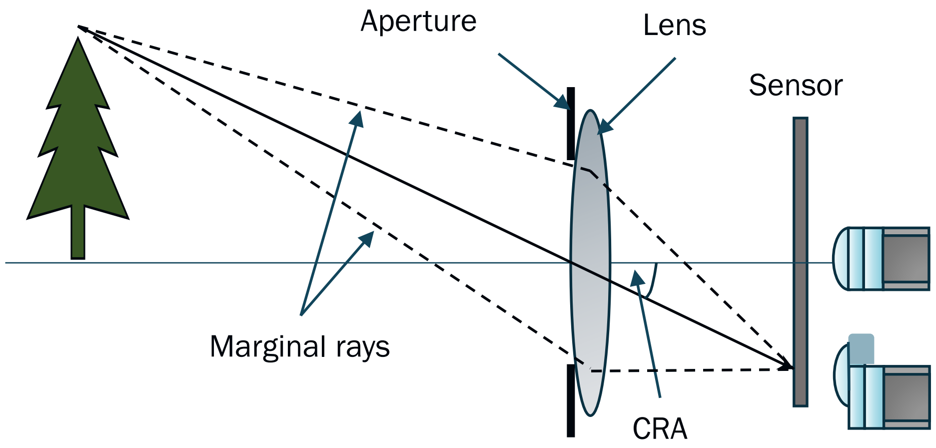






Chief Ray Angle (CRA)

- The angle of light rays that pass through the center of the exit pupil of the lens relative to the optical axis
- Often quoted as a number (for the maximum CRA), but it's actually a curve
- The microlenses on a sensor are (typically) shifted to compensate for the anticipated change in CRA across the array
- Mismatched lens and sensor CRA can lead to luma and chroma shading
  - Match for 1µm pitch pixels should be within ~2°
  - Match for 2µm pitch pixels should be within ~5°



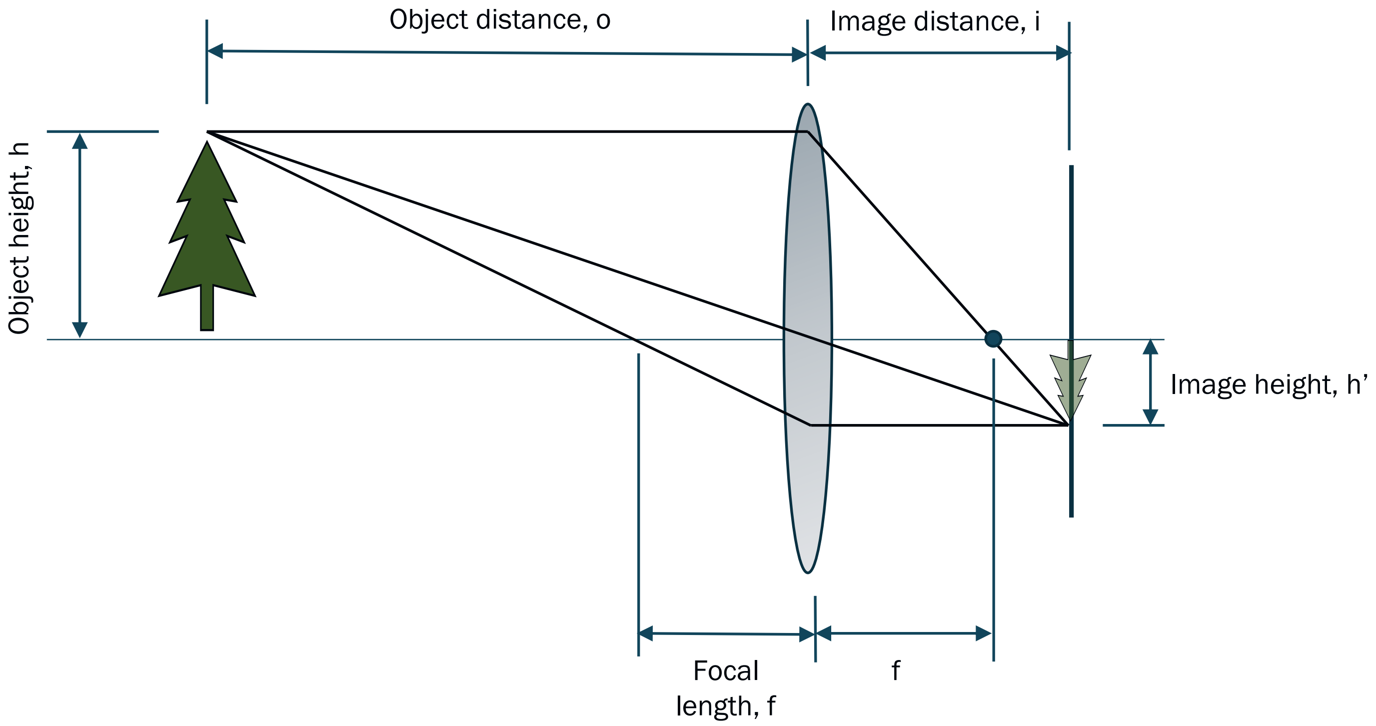
Focal Length & Magnification

- Focal length is a measure of how strongly the system converges (or diverges) light
- Thin lens approximation:

$$\frac{1}{o} + \frac{1}{i} = \frac{1}{f}$$

- Magnification, M:

$$M = \frac{-i}{o} = \frac{h'}{h}$$



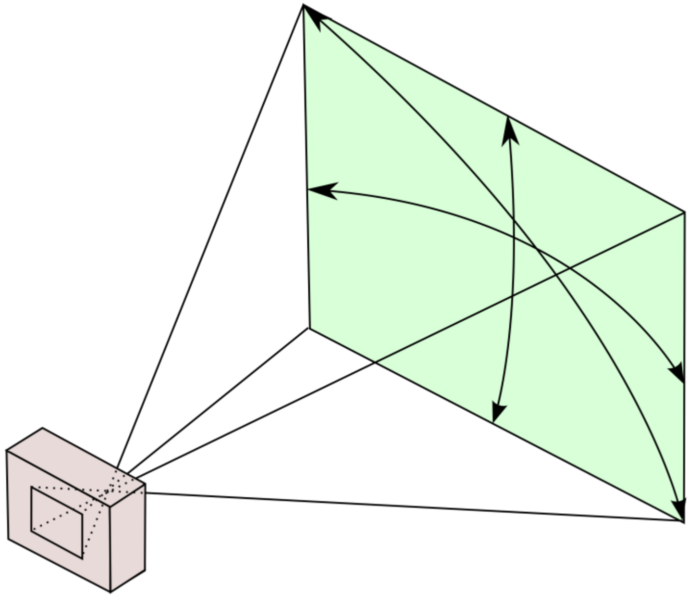
Field of View (FOV)

- FOV is determined by focal length and distortion
- For a rectilinear lens (one where straight features form straight lines on the sensor) the field of view, α, is:

$$\alpha = 2\arctan\frac{d}{2f}$$

Where d = sensor size, and f = focal length

- This is true so long as the object distance >> the lens focal length
- For lenses with non-rectilinear distortion (like GoPro lenses) the calculation gets more complex
- FOV is quoted as the diagonal FOV (DFOV), horizontal FOV (HFOV), or vertical FOV (VFOV)
  - Sometimes without defining which is being used



By Dicklyon at English Wikipedia - Transferred from en.wikipedia to Commons., Public Domain, https://commons.wikimedia.org/w/index.php?curid=10783200

F-Number

- Ratio of focal length to the diameter of the entrance pupil
  - Entrance pupil is the image of the physical aperture, as “seen” from the front (the object side) of a lens

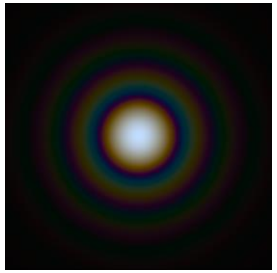
$$N = \frac{f}{D}$$

- The smaller N, the more light reaches the sensor
- A lens with a smaller f-number is often referred to as a “faster” lens, because it permits a faster shutter speed
- ”Stop” scale for f-number with each stop being 1/2 the amount of light passing through the optical system:

N	1.0	1.4	2	2.8	4	5.6	8	11	16	22	32

F-Number & Resolution Limits

- Diffraction of light by the aperture creates a lower “diffraction limit” to the resolution of the camera
  - This is another quantum mechanical effect
- A circular aperture creates an “Airy disc” pattern of light on the image sensor
- The smaller the aperture, the larger the airy disc
- For small pixels, or lenses with high f-numbers, real world performance can be limited by the diffraction limit
- Diameter of Airy disc to first null, d:
  - $d = 2.44\lambda N$
  - Where λ = wavelength and N f-number
- Table on right shows minimum resolvable pixel size vs f-number
  - This is for edge detection. For imaging points (like stars) the limits are higher



Simulated Airy Disc for white light

By SiriusB - Own work, CC0, https://commons.wikimedia.org/w/index.php?curid=68302545

F-Number	Blue (430nm)	Green (550nm)	Red (620nm)
1.0	0.44 µm	0.54 µm	0.61 µm
1.2	0.53 µm	0.64 µm	0.73 µm
1.4	0.61 µm	0.75 µm	0.85 µm
1.6	0.70 µm	0.86 µm	0.97 µm
1.8	0.79 µm	0.97 µm	1.09 µm
2	0.88 µm	1.07 µm	1.21 µm
2.2	0.97 µm	1.18 µm	1.33 µm
2.4	1.05 µm	1.29 µm	1.45 µm
2.6	1.14 µm	1.40 µm	1.57 µm
2.8	1.23 µm	1.50 µm	1.69 µm
3	1.32 µm	1.61 µm	1.82 µm
3.2	1.41 µm	1.72 µm	1.94 µm
3.4	1.49 µm	1.83 µm	2.06 µm
3.6	1.58 µm	1.93 µm	2.18 µm
3.8	1.67 µm	2.04 µm	2.30 µm
4	1.76 µm	2.15 µm	2.42 µm

Depth of Field (DoF)

- Depth of field is the distance between the nearest and farthest objects that are acceptably in focus

$$DoF \approx \frac{2o^2\mathrm{Nc}}{f^2}$$

- Where o = object distance, N = f-number, c = circle of confusion (acceptable blur circle), f = focal length
- DoF increases with:
  - Focus distance (squared)
  - Circle of confusion
  - F-number
    - At the cost of light reaching the sensor, and the diffraction limit
- DoF decreases with:
  - Focal length (squared)
    - A shorter focal length lens has a larger depth of field. So either a wider FOV, or a smaller sensor will increase the DoF

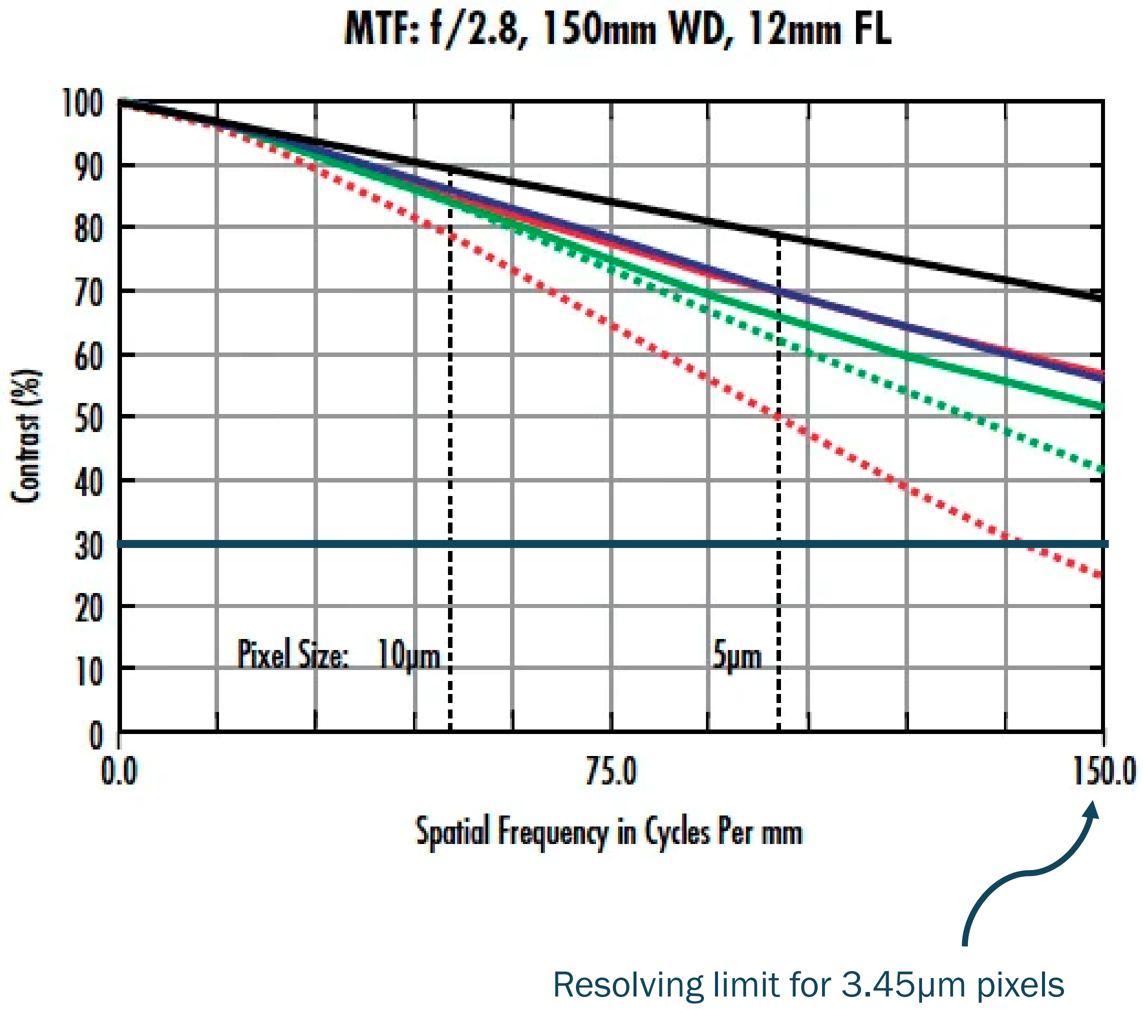
Wavelength Range

- A lens is designed for a specific wavelength range
- The wider the wavelength range the more difficult the design becomes
  - Leading to higher cost/performance ratio
- A lens designed for visible light will have poor performance in the NIR, and visa versa
- So a lens should be selected to match the application
- A camera for NIR only, should use a NIR lens
- If using visible + NIR, a lens designed for the full wavelength range should be selected

Modulation Transfer Function (MTF)

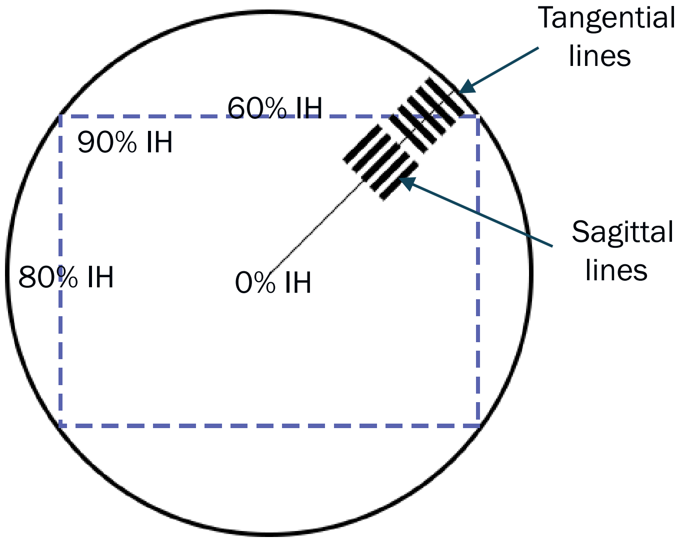
- MTF is a measurement of the sharpness and contrast of a lens over a range of spatial frequencies (starting at 0)
- MTF plots allow a lens to be evaluated for a particular application, and lenses to be compared
- A MTF score of 30% means that an edge will be well resolved
- Horizontal axis is line pairs per mm. Higher lp/mm -> smaller features
- The highest spatial frequency that can be sampled by the image sensor is:

$$\frac{lp}{mm} = \frac{1000\,\frac{\mu m}{mm}}{2\ x\ Pixel\ Size\ (\mu m)}$$



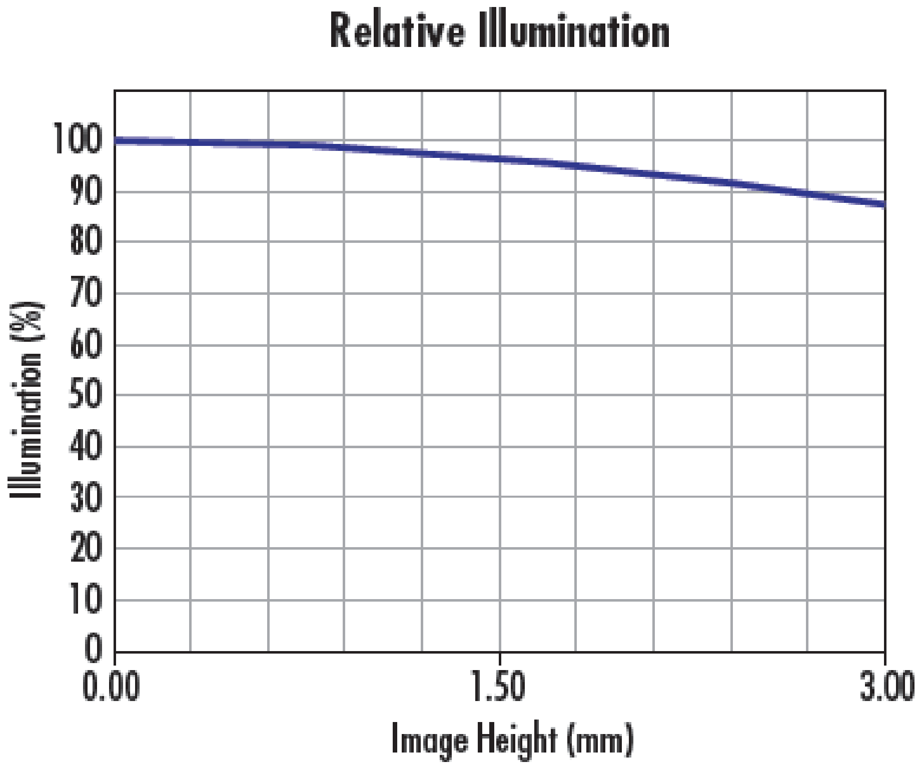
MTF (Continued)

- This type of MTF plot typically contains lines for various image heights
  - For example 0% IH, 60%IH, 90% IH
  - This allows MTF in the center, middle, and corners of the image to be assessed
- Plots typically show different lines for sagittal and tangential directions
- Alternative plots show MTF vs image height for specific spatial frequencies
- Warning on MTF:
  - Unless stated, quoted MTF is for an “as designed” lens, without manufacturing tolerances
  - Actual ”as built” MTF will always be worse that this
    - This can be simulated using Monte Carlo simulations, and measured on optical test equipment



Relative Illumination (RI)

- Relative illumination (RI) is a measure of the roll-off in light intensity from the center to the edge of the lens image circle
- Often stated as a single number, being the RI at 100% IH
- Plot on right is for a lens with a 88% RI
  - This is very good, a typical mobile phone lens, by contrast, will have an RI <30%
- This shading effect means corners will be noisier (lower SNR) than the center of the image
- The sensor will also have some roll-off in sensitivity with CRA, so camera RI is the product of the lens RI and the sensor roll off
- RI can be corrected by applying a spatially-varying normalization function, but this amplifies noise in the corners



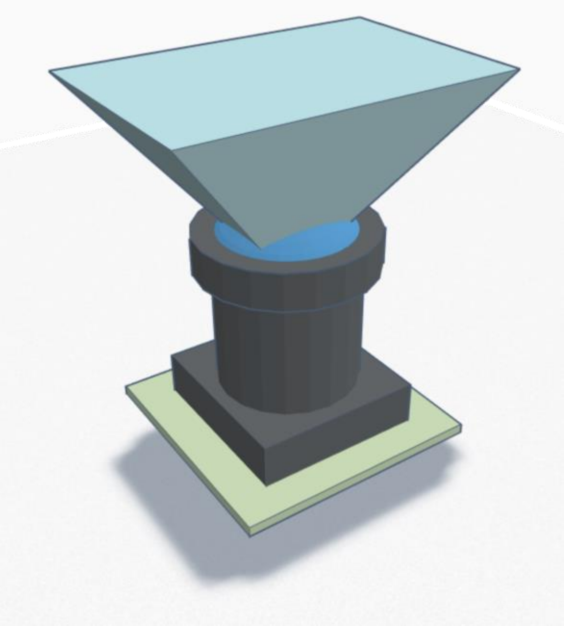
Stray Light & Ghosts

- Stray light is light scattered within the optical system that reduces contrast
- Ghosts are spatially localized reflections that can obscure details or confuse machine vision
- Sources include:
  - Imperfect anti-reflective coatings
    - Typical coatings on glass lens elements reflect <0.3% of normal incident light
    - Typical coating on plastic lens elements reflect <0.4% of normal incident light
  - Edges of lens elements (sometimes inking helps)
  - Mechanical surfaces (can be blackened or textured)
  - Diffraction off the image sensor
  - Dirt and oils on optical surfaces
- Can be simulated using the full lens design and specialist tools (Fred, LightTools, ASAP)
- Can be measured in the lab using a point light source (fiber bundle) and a robotic stage to rotate the camera

Faux stray light rendered in Adobe Photoshop. Subjectively good for art, bad for object recognition

Vignetting

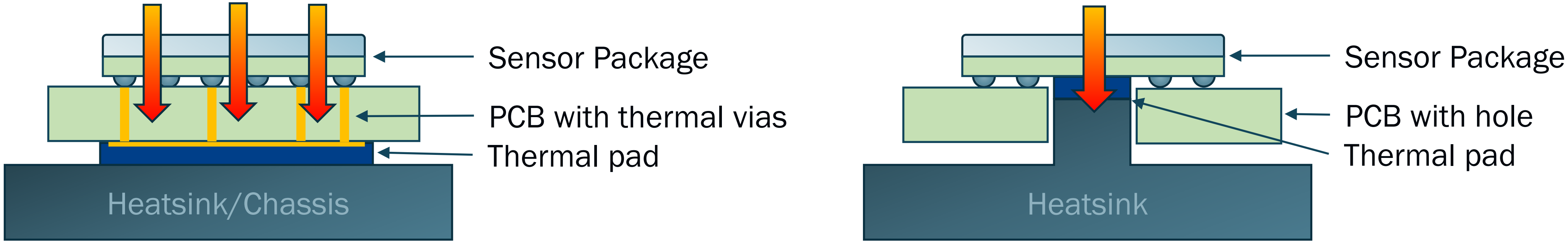
- Vignetting is unintended blocking of light rays towards the edge of the image
- Check all mechanical features for possible blocking
  - Filter/window holders
  - Main product housing
- Have an optical engineer generate a mechanical FOV for importing in to CAD
- Include mechanical tolerances in determining clear window sizing
  - Lens to sensor centration
  - Lens to sensor tip/tilt
  - Camera module to housing alignment tolerances (x, y, z, tip/tilt)



Other Design Factors

Thermal Considerations

- Dark current doubles about every 6°C
- Dark signal non-uniformity increases with temperature
- Dark signal shot noise (= SQRT number of dark signal electrons) can become a factor at high temperatures
- Hot pixels increase with temperature (defect thermal activation)
- Black level corrections get worse with temperature
- Thermal design of a camera should be considered *up front*
- Design a low thermal resistance path from the sensor die to ambient



Electrical Factors (Keep it Quiet)

- Image sensors are mixed signal devices
- Care must be taken to ensure low power supply noise, particularly on analog supply rails
  - Dedicated supply (LDO) for sensor V_{ANA} is a good idea
- Careful placement and low-impedance routing of decoupling capacitors is essential
  - Ask sensor supplier for a module design reference guide
  - Prioritize placement of analog rail decoupling capacitors (place as close to the sensor as possible) with minimal vias
- Analyze the noise
  - Capture dark raw images (no image processing) and extract noise components
  - Analog supply noise is particularly prone to creating temporal row noise
    - Compare with sensor noise on manufacturers evaluation system, or other reference design

Selecting a Sensor & Lens

Selecting a Sensor - Determine Key Requirements

1. Determine resolution requirements:
   - Object space resolution
   - Field of view
   - Max. object distance
   - Min. linear pixels for feature recognition

$$Pixel\ Size\ (\mu m) = Object\ Space\ Resolution\ (\mu m) \times \frac{FOV\ (mm)}{Sensor\ Size\ (mm)}$$

2. Wavelength of light to be imaged: visible or NIR-enhanced sensor
3. Color (full RGB, mono + R, etc.), or monochrome
4. Dynamic range requirements
5. High-speed motion (camera or scene): rolling/global shutter
6. Interface requirements for SOC

Selecting a Sensor – Evaluate using Sensor Evaluation Kit

- Use requirements from previous slide to select a sensor evaluation kit (EVK)
  - Look for sensors intended for similar applications
  - Many sensor companies have online product selection tools, or can support at the sales/application engineering level
- If possible, pair EVK with lens that's a close fit for final application (see following section)
- Test EVK in your application including with minimum illumination conditions
  - If lens FOV does not match, adjust object distance, or chart size to compensate
- If low light SNR is acceptable proceed to lens selection
- If low light SNR is not acceptable:
  - Can required SNR be achieved by adjusting the lens aperture?
    - Trade-off = depth of field, lens size/cost
  - Can required SNR be achieved by increasing exposure time?
    - Trade-off = motion blur
  - If SNR requirements cannot be achieved, select a larger sensor (more light collection) and repeat process
    - Trade-off = depth of field, sensor cost, lens size/cost
- Evaluate dynamic range
  - This is one of the more difficult tests to perform correctly, but online resources exist like Imatest test software tutorials

Selecting a Lens

1. Determine optical format (from sensor size)
2. Select field of view ⬄ focal length
3. Determine target f-number
4. Check depth of field
5. MTF
   - From resolution study, determine line pairs/mm resolution needed
   - An MTF ≥ 30% at the target lp/mm is a good rule of thumb for the detection of edges
   - Check MTF across field: center, mid, corner regions
6. Pick a lens with a CRA that closely matches the sensor CRA specification
7. Check the impact of relative illumination on performance in corners (SNR decrease)
8. Determine if the lens includes a filter (IR-cut, IR passband, etc.), if one is needed
9. Check if the lens is suitable for the dynamic range requirements (HDR requires very good anti-reflective coatings)

Resources

Image Sensor Manufacturers

ON Semiconductor: ON Semi Image Sensors

Sony: Sony image sensors

Omnivision: https://www.ovt.com/

Samsung: Samsung image sensors

Fairchild Imaging (BAE Systems):

https://www.fairchildimaging.com/

ams: https://ams.com/cmos-imaging-sensors

Interfaces, Optics, and Imaging

MIPI Alliance: https://www.mipi.org/

Edmunds Optics Knowledge Center:

- Optics: https://www.edmundoptics.com/knowledge-center/#!&CategoryId=114
- Imaging: https://www.edmundoptics.com/knowledge-center/#!&CategoryId=175

Cambridge in Colour imaging tutorials:

https://www.cambridgeincolour.com/tutorials.htm

Imaging Test Software and Charts

Imatest: https://www.imatest.com/