Selecting the Right Camera for Your Embedded Computer Vision Project

Adrian Marques
Managing Partner
Digital Sense

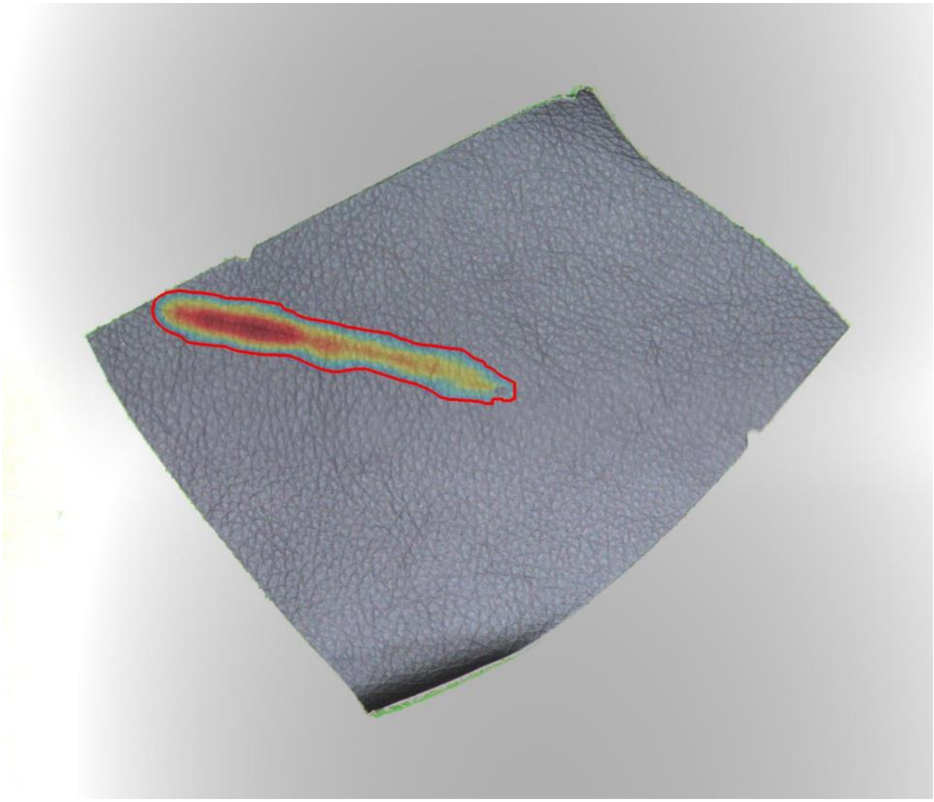
Digital Sense

- **Computer Vision & Machine Learning Studio**
- 100+ successful projects, working with startups and Fortune 500 companies
  - Agritech
  - Geospatial
  - Industrial inspection, and more
- **Highly-qualified team:**
  - 7 PhDs, 8 Masters, 200+ publications, 7 patents
- **Focused on R&D and innovation**

Visit us at digitalsense.ai

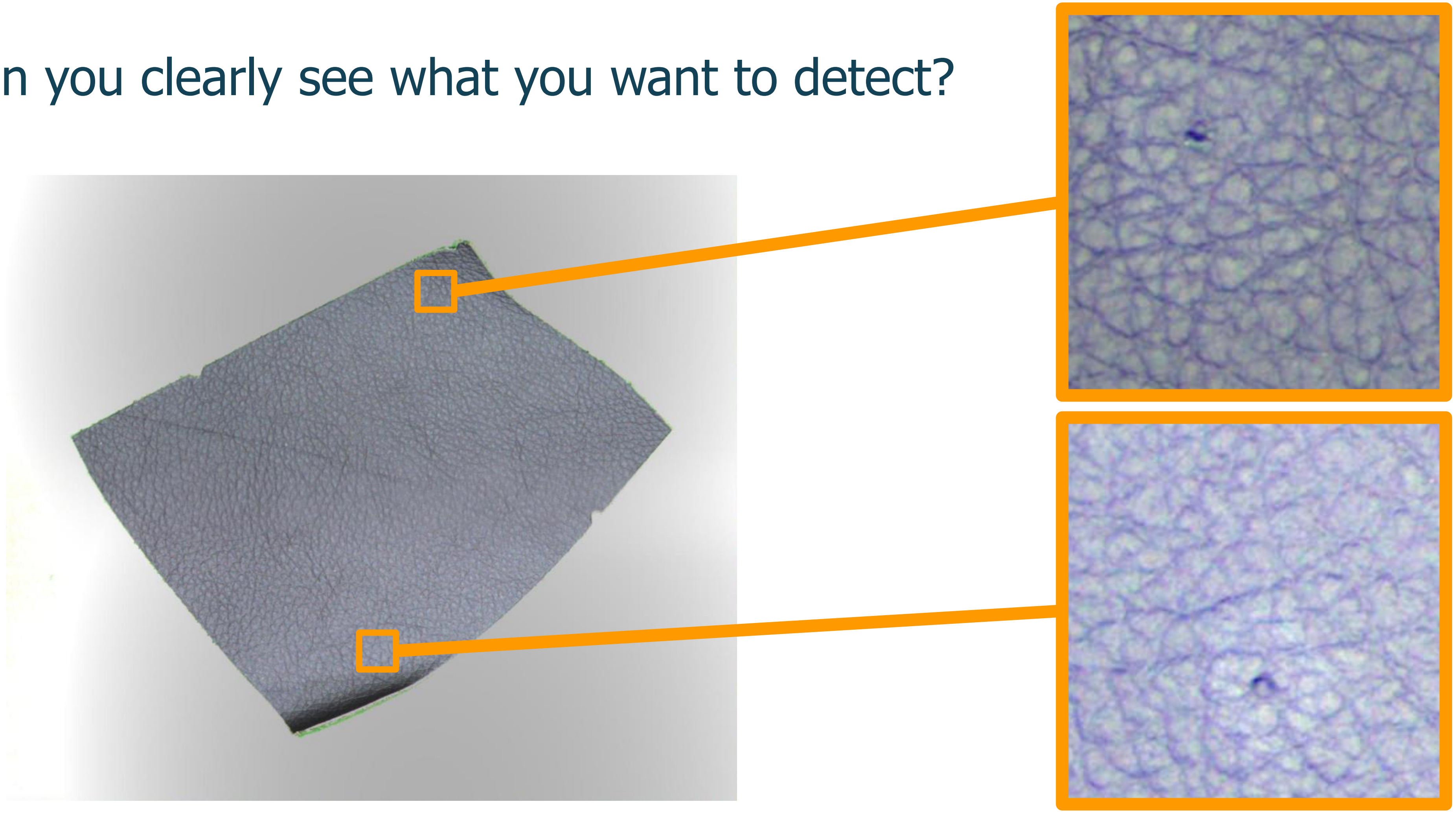
Motivation

- A well-designed acquisition system can be key to the success of a computer vision project



Motivation

- Can you clearly see what you want to detect?



Motivation

- Is your input sharp, highly contrasted and noiseless?

CODE READ OK

CODE READ FAILED

CODE READ FAILED

Motivation

- Are you losing important information in your input?

Visible spectrum camera UV camera

Image: Birchbox

Motivation

- To ensure **consistent and robust performance** in industrial applications, **the design of the acquisition system is fundamental**.
- There are many aspects to consider that may seem confusing or overwhelming when first facing them.
- In this presentation we'll introduce some of the most common elements you'll need to take into account when choosing your camera.

Some questions

- What wavelengths are your objects of interest most reflective/absorbent in?

  Spectral bands

- What's their size?
- What's their distance to the camera?

  Resolution, field of view

- How fast are they moving relative to the camera?
- Do you expect your camera to vibrate/shake?

  FPS, global vs rolling shutter

Some questions

- What are the lighting conditions like?
  - Indoors or outdoors?
  - Controlled or natural lighting?
  - Low light setting?

    Dynamic range, HDR

- Do you need to conform to a form factor or interface? How far is your processor from your camera?

    Interfaces

- How rough is your operating environment? What's your budget? What level of support will you provide for your application?

Outline

- Image formation basics
- Spectral bands
- Sensors, continued
- Lenses
- Configuration
- Interfaces & other considerations
- Conclusion and recommendations

Image formation basics

DigitalSense

Image formation basics

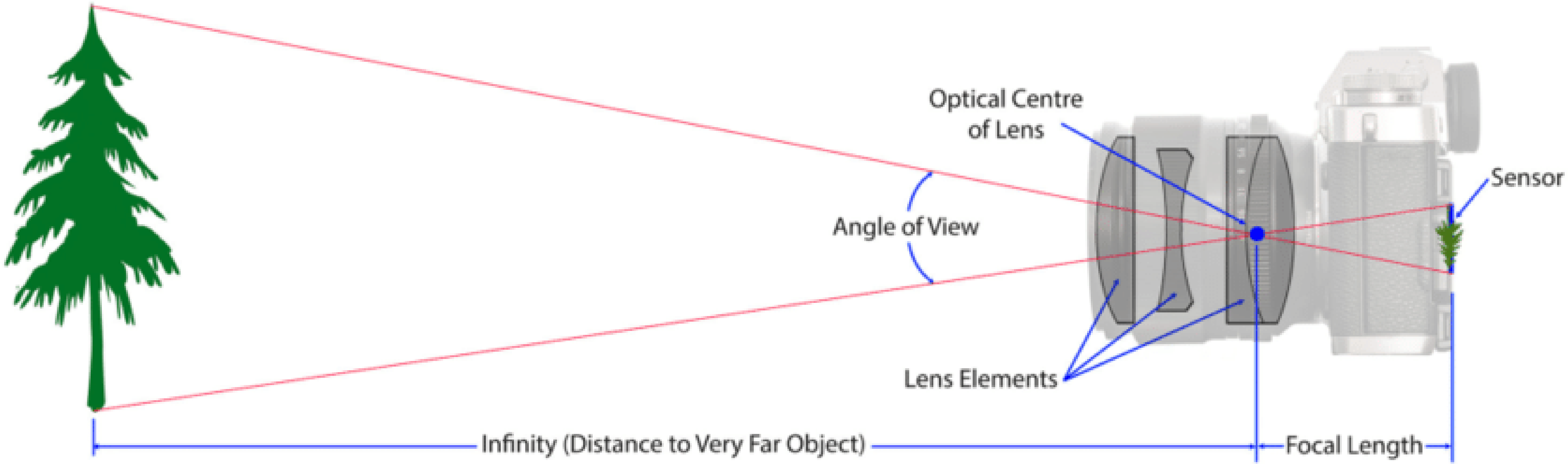
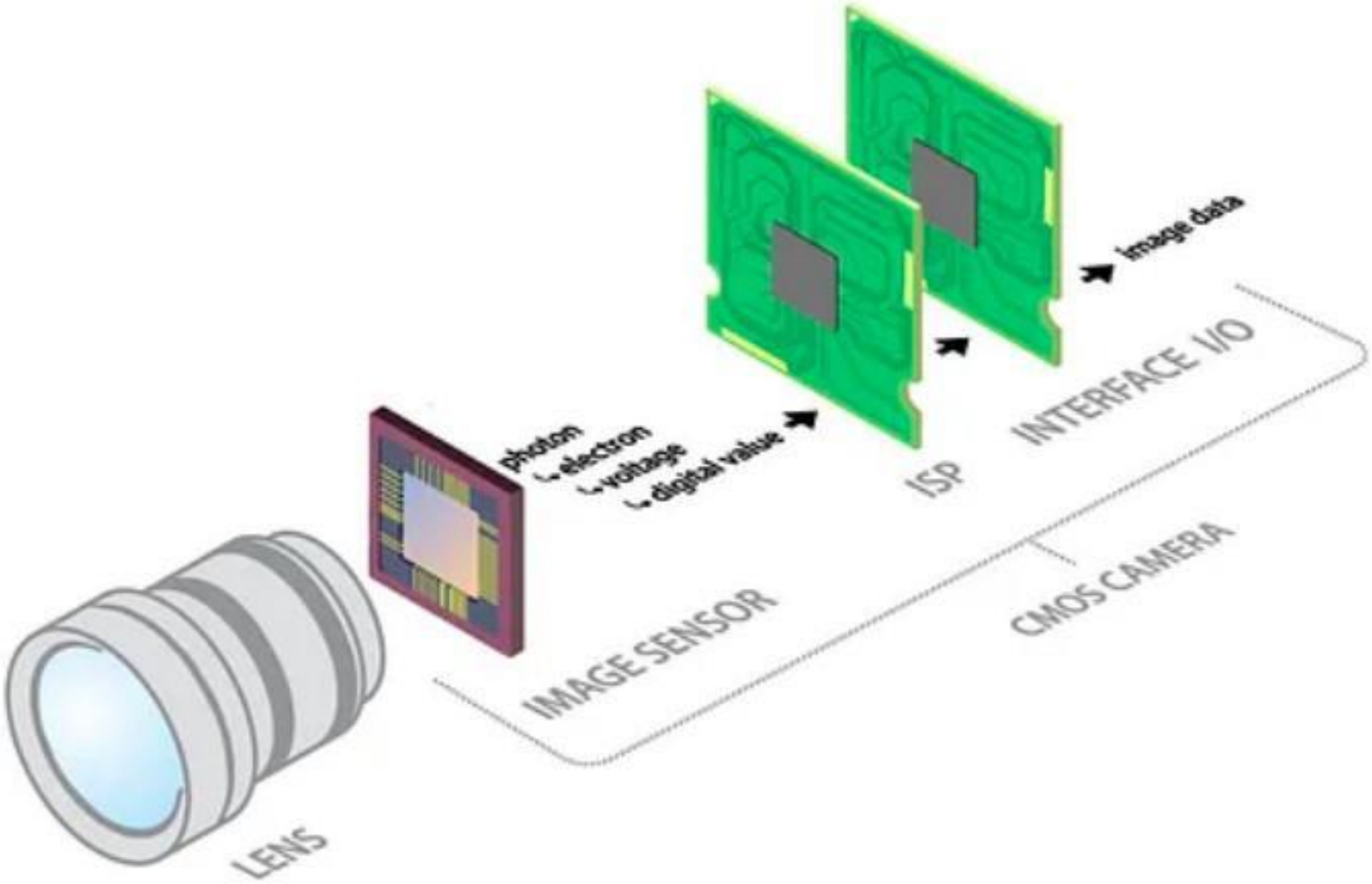
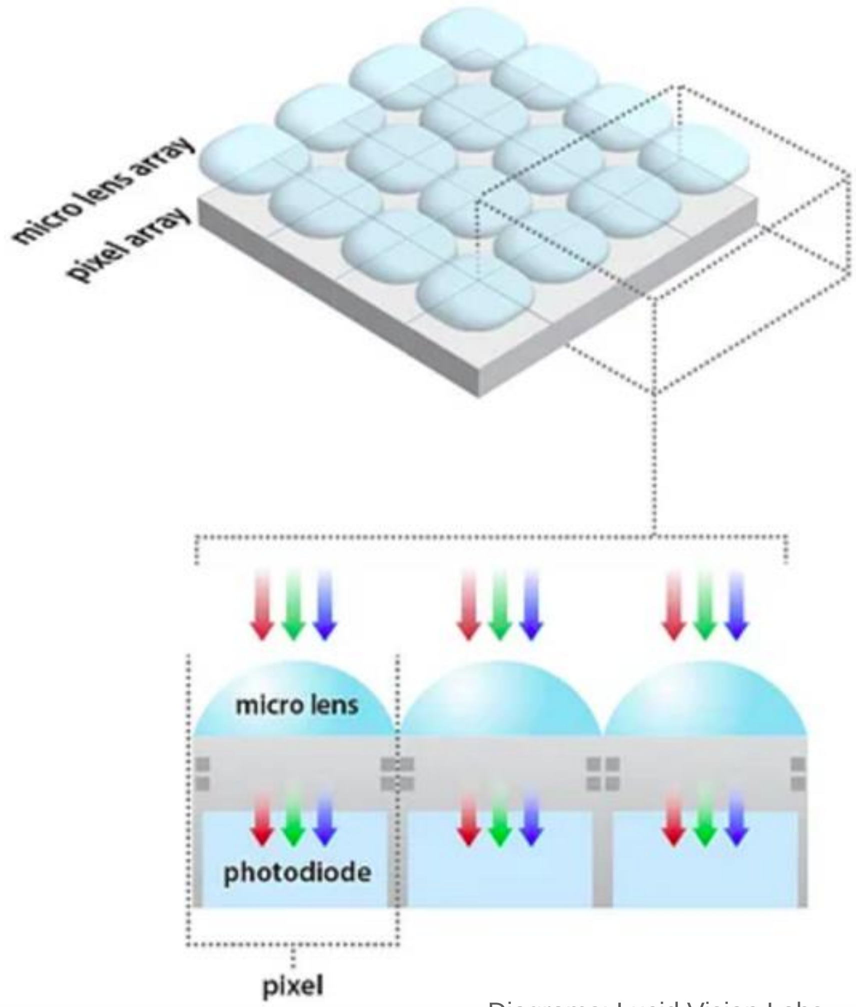


Diagram: Cemal Melih Tanis

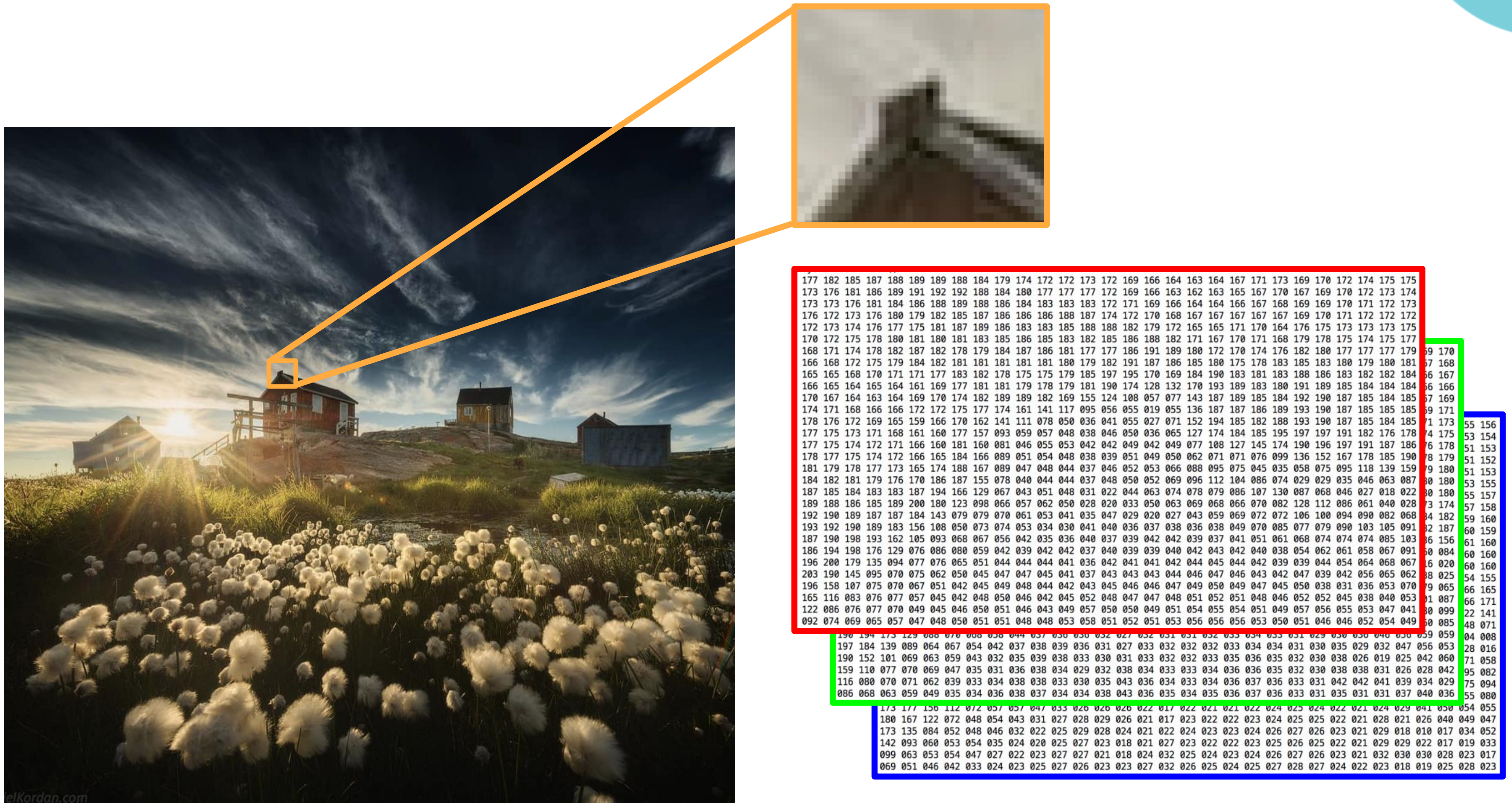
Image formation basics





Diagrams: Lucid Vision Labs

Image formation basics

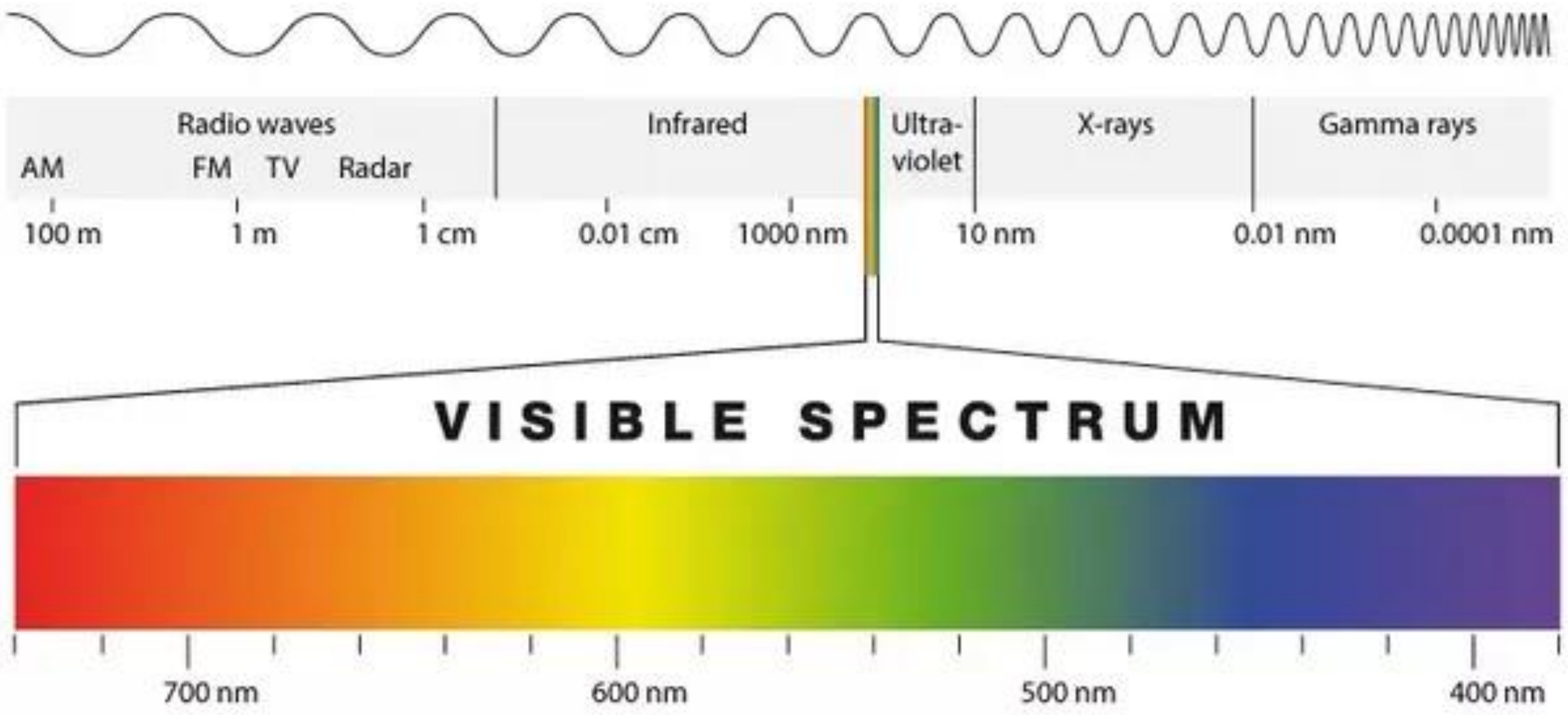


Spectral bands

DigitalSense

Spectral bands

- Pick your bands:
  - Monochrome
  - RGB
  - Infrared
  - Multispectral
  - Thermal
  - Custom filters



Visible spectrum range: 750 nm (red) to 380 nm (violet)

Diagram: Analytic

Spectral bands

- **RGB:**
  - Mostly realized by having R, G or B filters over individual pixels in a mosaic.
  - Full image is reconstructed from these lower resolution samples.
- **Monochrome:**
  - higher resolution and light sensitivity than RGB.
- **Shortwave infrared, thermal:**
  - Pixels are constructed differently than for RGB cameras.
  - Lower resolution and are more expensive than RGB.

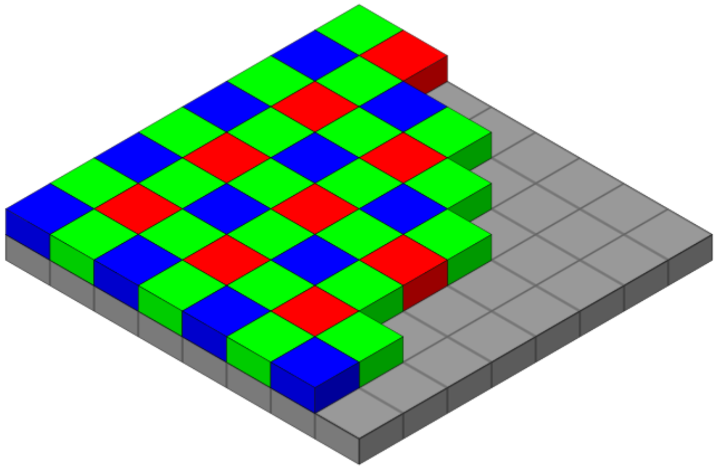


Diagram: Cburnett

Spectral bands examples

Optical (380 to about 750 nanometers)

Thermal(8 to 14 micrometers)

Spectral bands examples

Typical Spectral Reflectance Curve for Vegetation

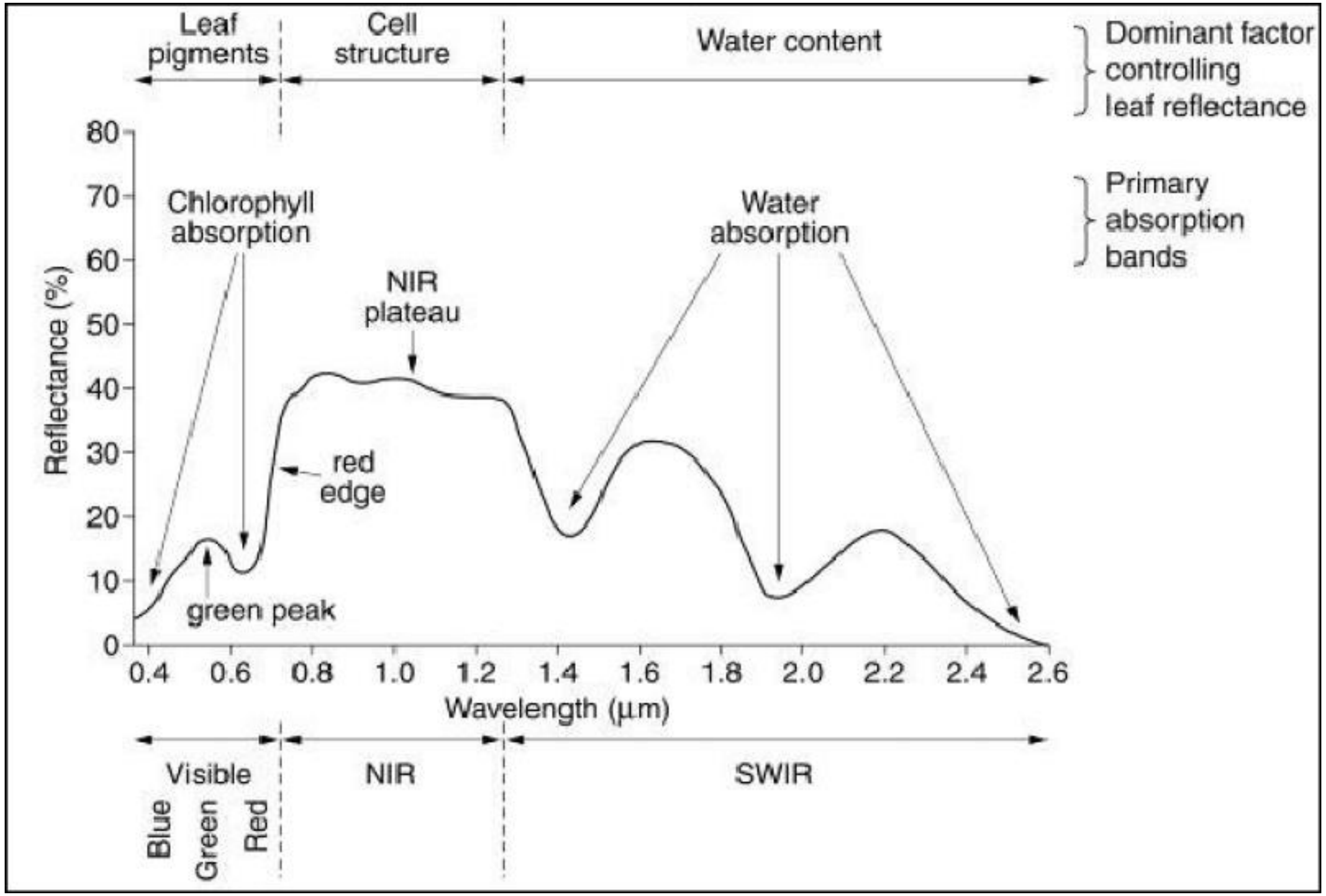
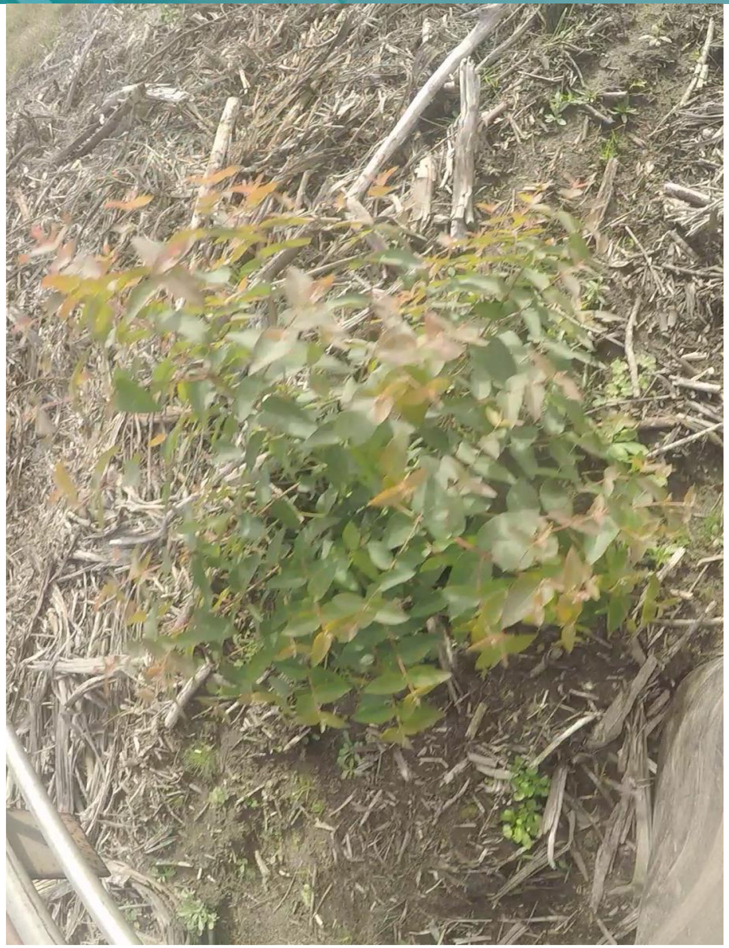


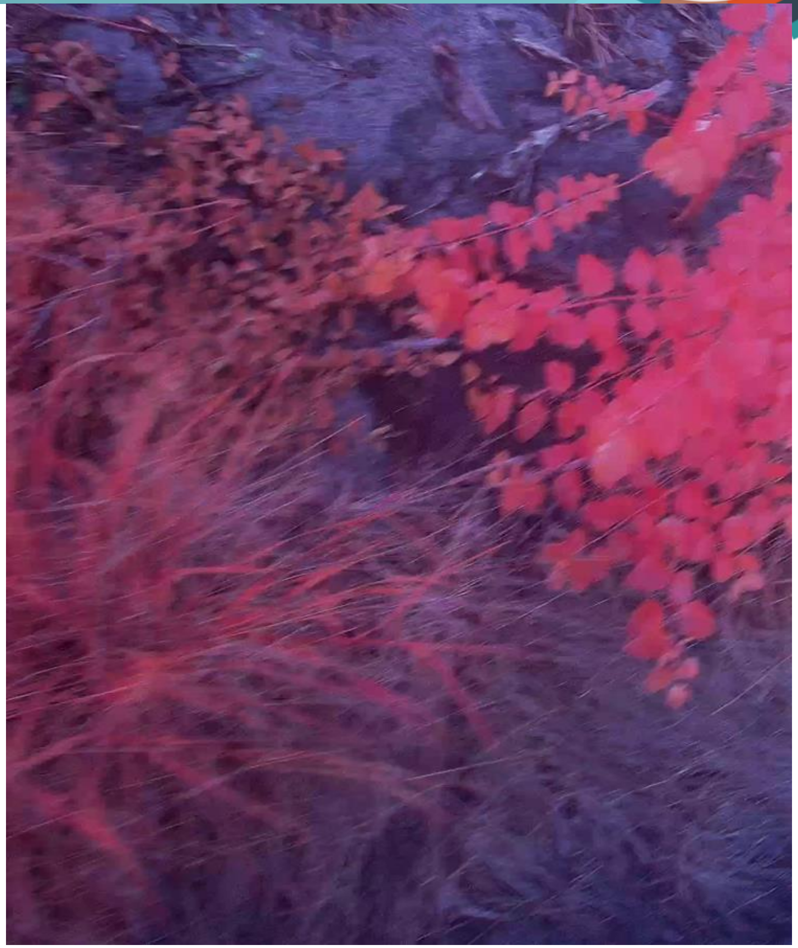
Image via Akhil Kallepalli

Spectral bands examples



RGB

RGN

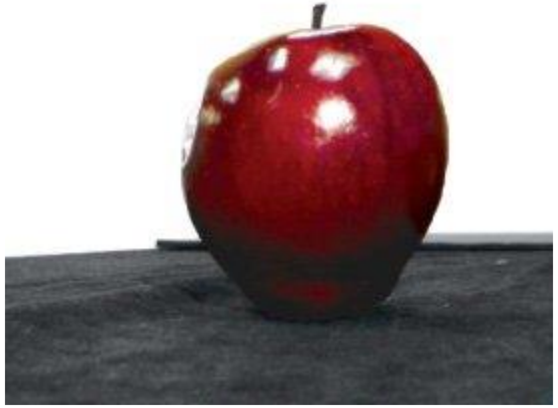


NGB

DigitalSense

Spectral bands examples: RGB vs SWIR

SWIR: short-wave infrared



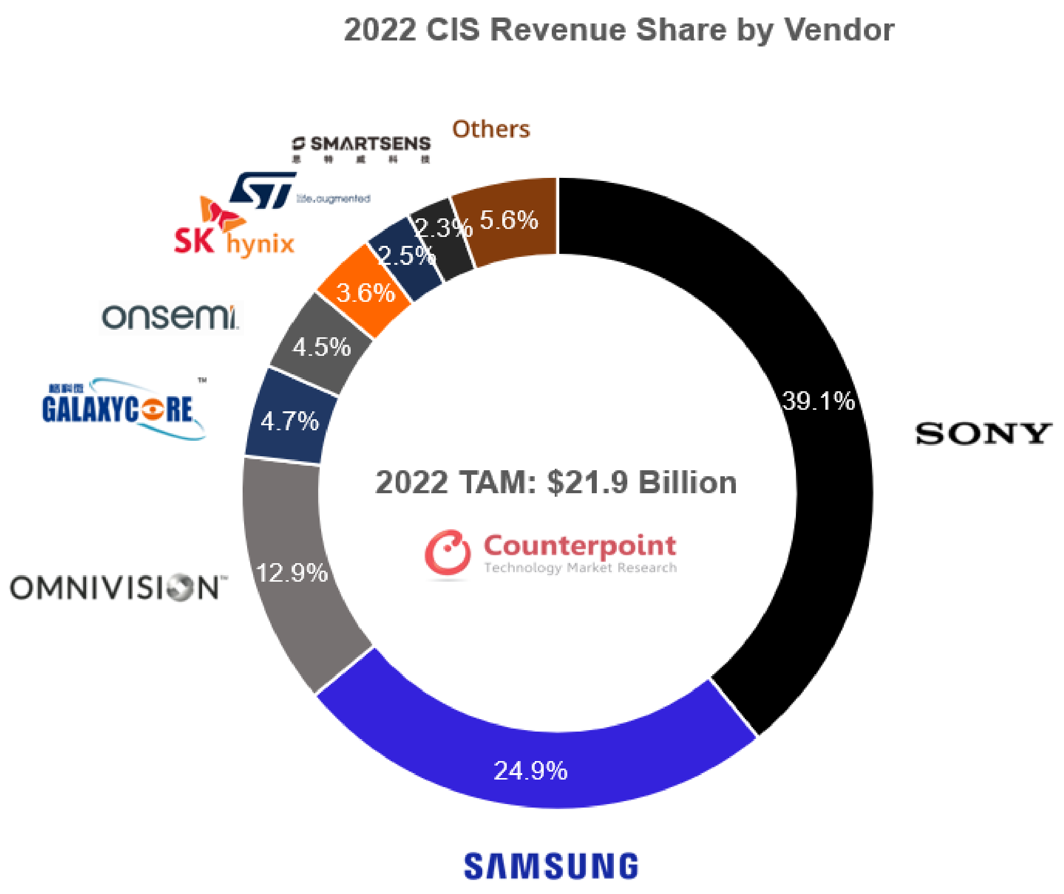
Bruising becomes evident under SWIR imaging.

Some materials can be opaque in RGB but transparent in SWIR.

Images: Edmund Optics

Sensors, continued

Sensor manufacturers



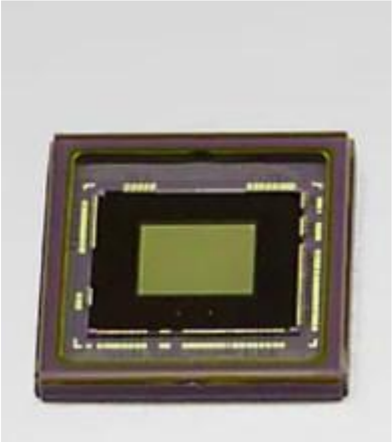
Source: Counterpoint Technology Market Research

- When choosing a camera, the sensor will come from one among a handful of manufacturing companies.
- Two cameras with the same sensor may have different imaging quality based on sensor integration and ISP features.

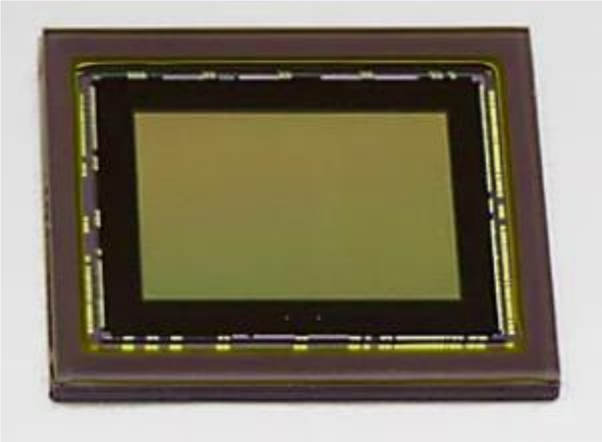
Sensor and pixel size

- Sensors come in different size formats.
  - Number of pixels in the sensor and pixel size determine sensor size.

1/6"



1/3"



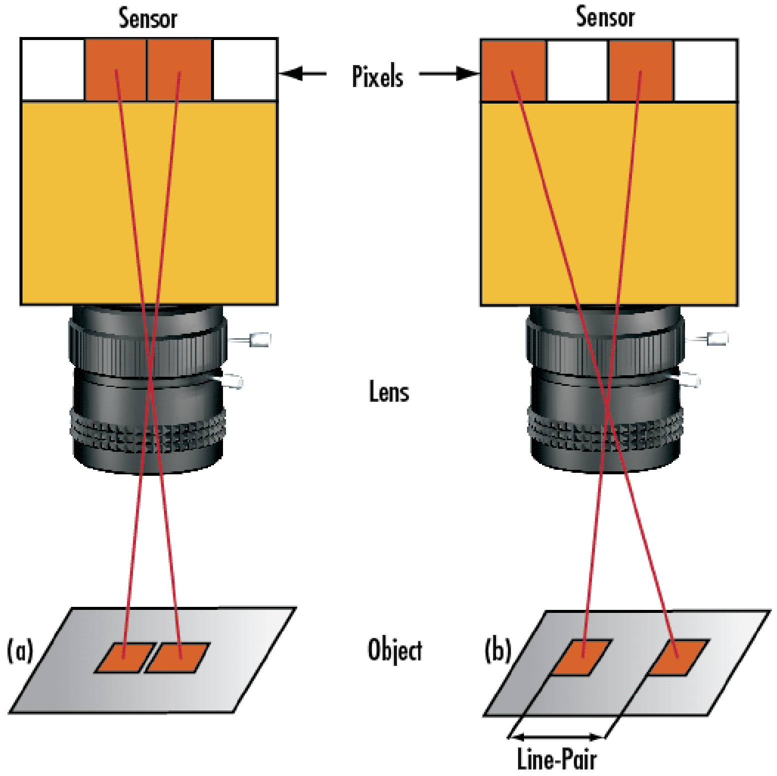
1"

Images: Lucid Vision Labs

- Typical pixel size range: 1 to 9 µm
  - Larger pixels capture more photons, leading to better SNR and dynamic range.

Resolution

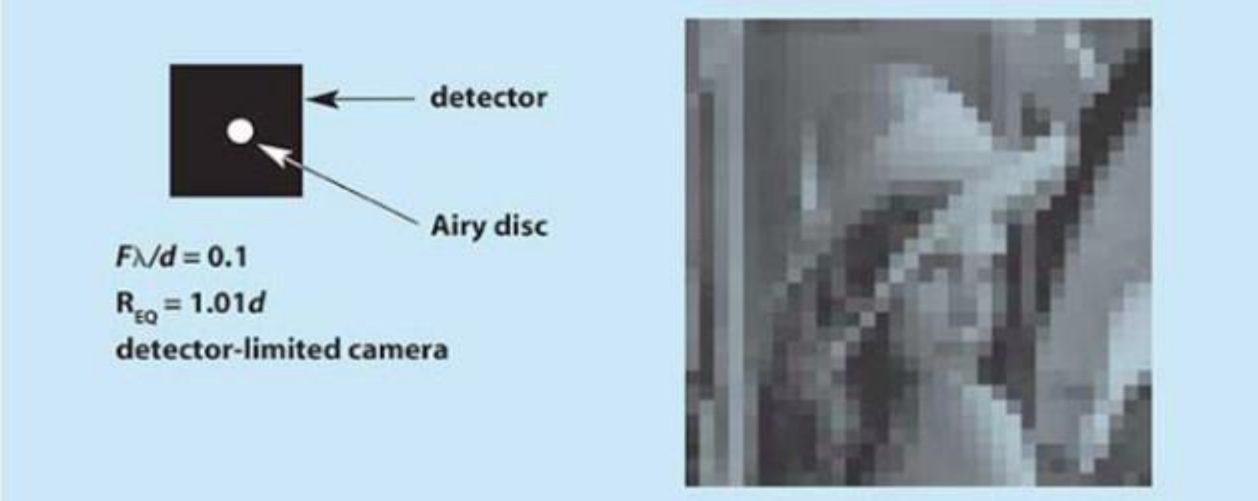
- The **sensor's resolution:** its number of pixels.
- The **camera's resolution** can be defined as the minimum distance between two distinguishable points in an image.



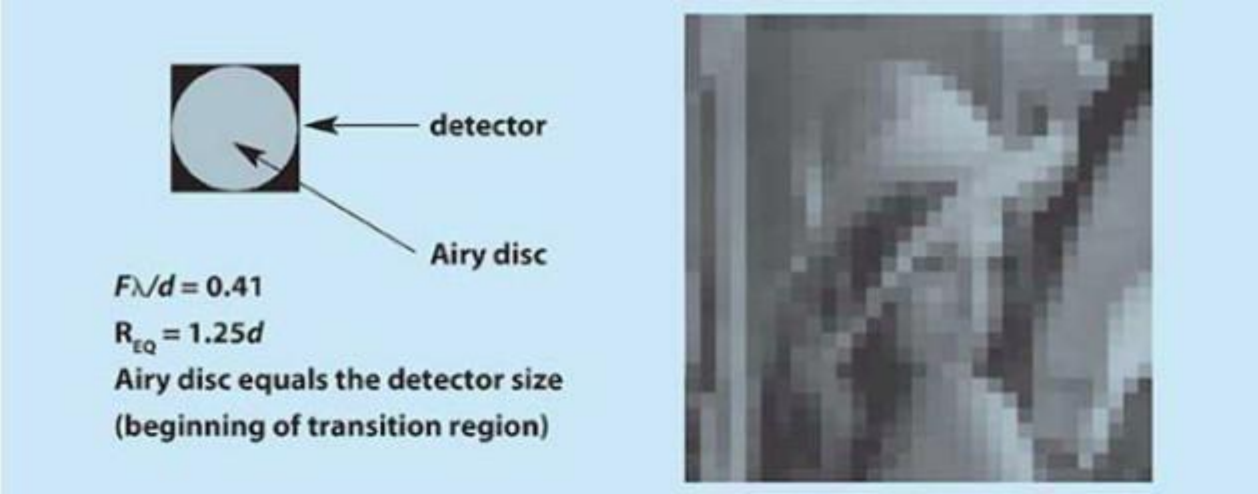
- It is strongly correlated with the number of pixels in the sensor but affected by other factors:
  - Presence of a bayer filter
  - Lens sharpness

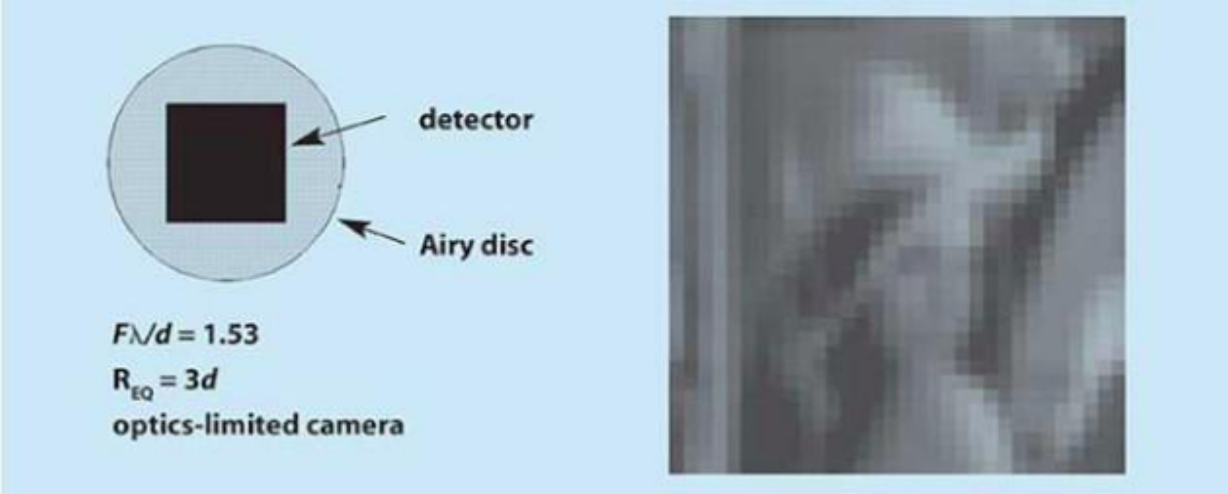
Diagram: Edmund Optics

Resolution



Lens sharpness and effect on image resolution.





Resolution

- What will be the actual **size in pixels of the objects of interest?**
  - Consider sensor resolution and field of view
- High **resolution can be wasted if you downscale** to a deep network's input size
- Some example standard input sizes for state-of-the-art architectures:
  - **Yolov5:** 640x640
  - **Resnets:** 224x224

Global vs rolling shutter

- **Global shutter:**
  - All lines exposed at once
- **Rolling shutter:**
  - Lines exposed sequentially
  - Introduces deformations in fast-moving objects
  - Can be an important issue if the camera vibrates or for these fast-moving objects

Images: Sony

Global vs rolling shutter

Global vs. rolling shutter in component inspection

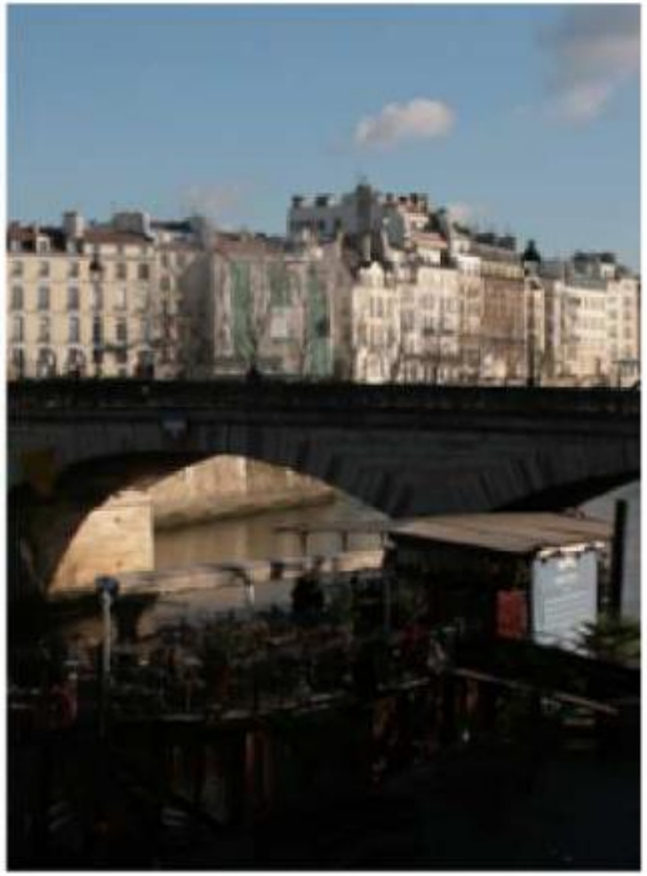
- Rolling shutter typically offers other advantages:
  - Higher frame rates
  - Lower noise
  - Lower cost

Images: Edmund Optics

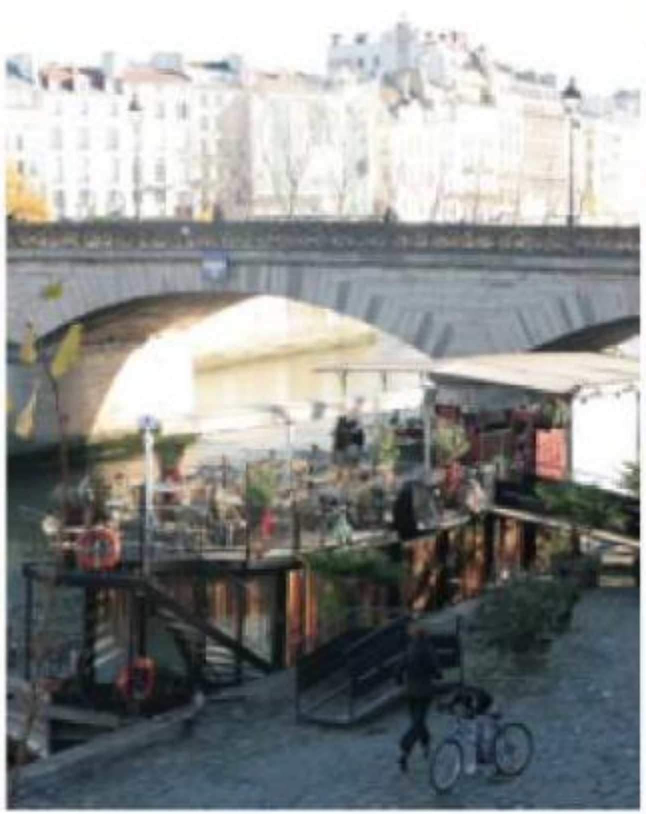
Dynamic range

Ratio between largest non-saturating input signal and the smallest detectable input signal

Natural scenes	100+ dB
Human eye	Around 90 dB
Cams with no HDR	70 dB
Cams with HDR	120+ dB



Lower exposure: dark areas are underexposed



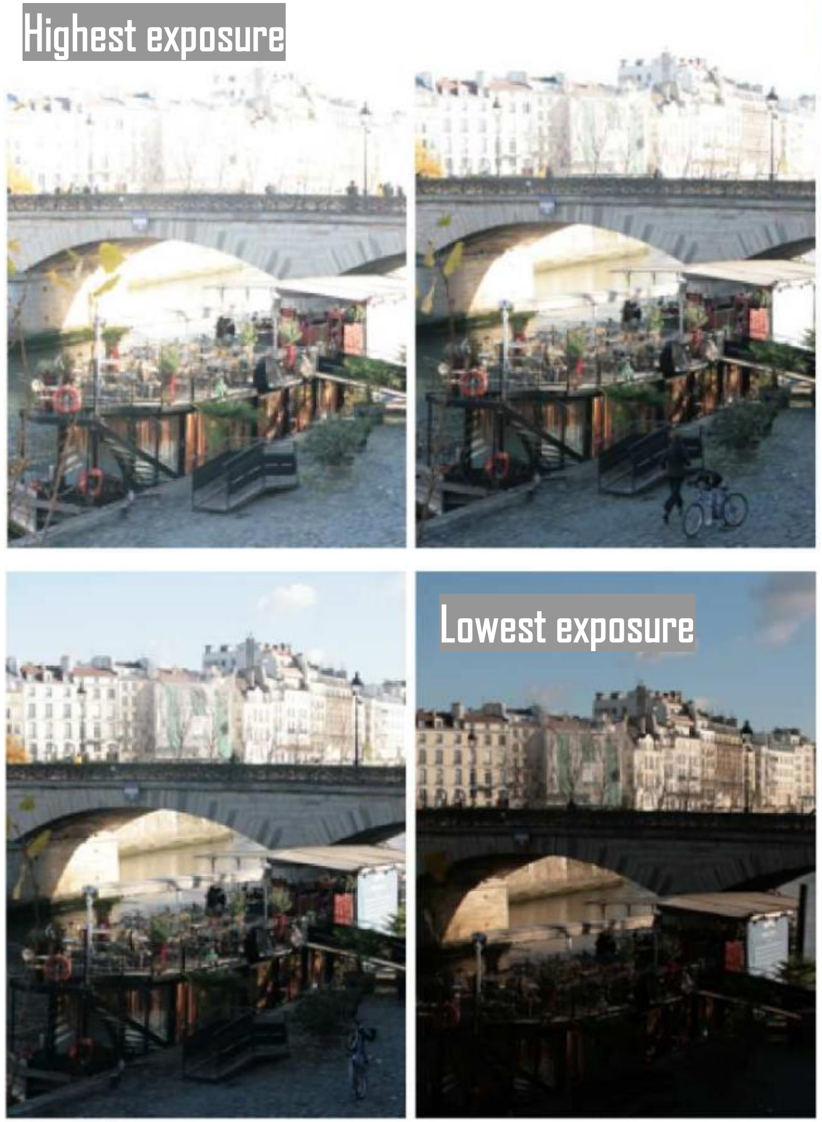
Higher exposure: bright areas are saturated

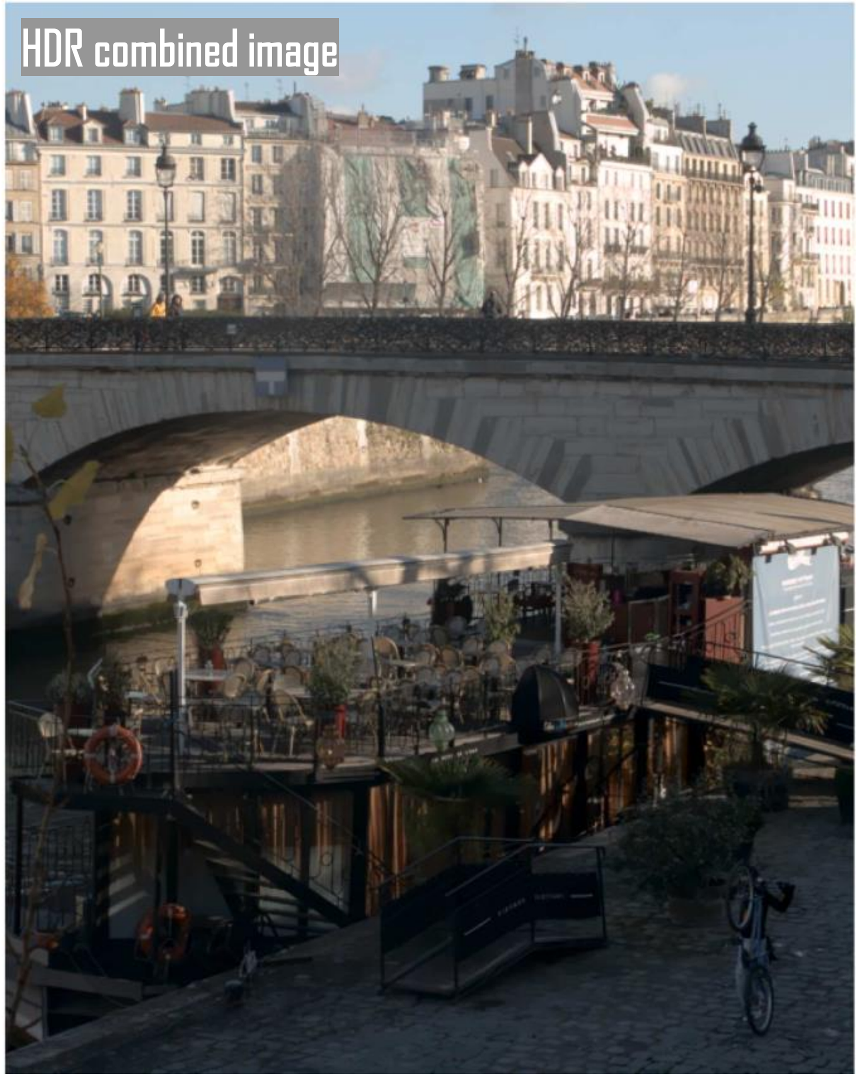
Images: Cecilia Aguerrebere et al.

HDR: multi-capture

Images at different exposures are combined into a single high-dynamic-range image

- Motion in HDR shots can result in ghosting artifacts

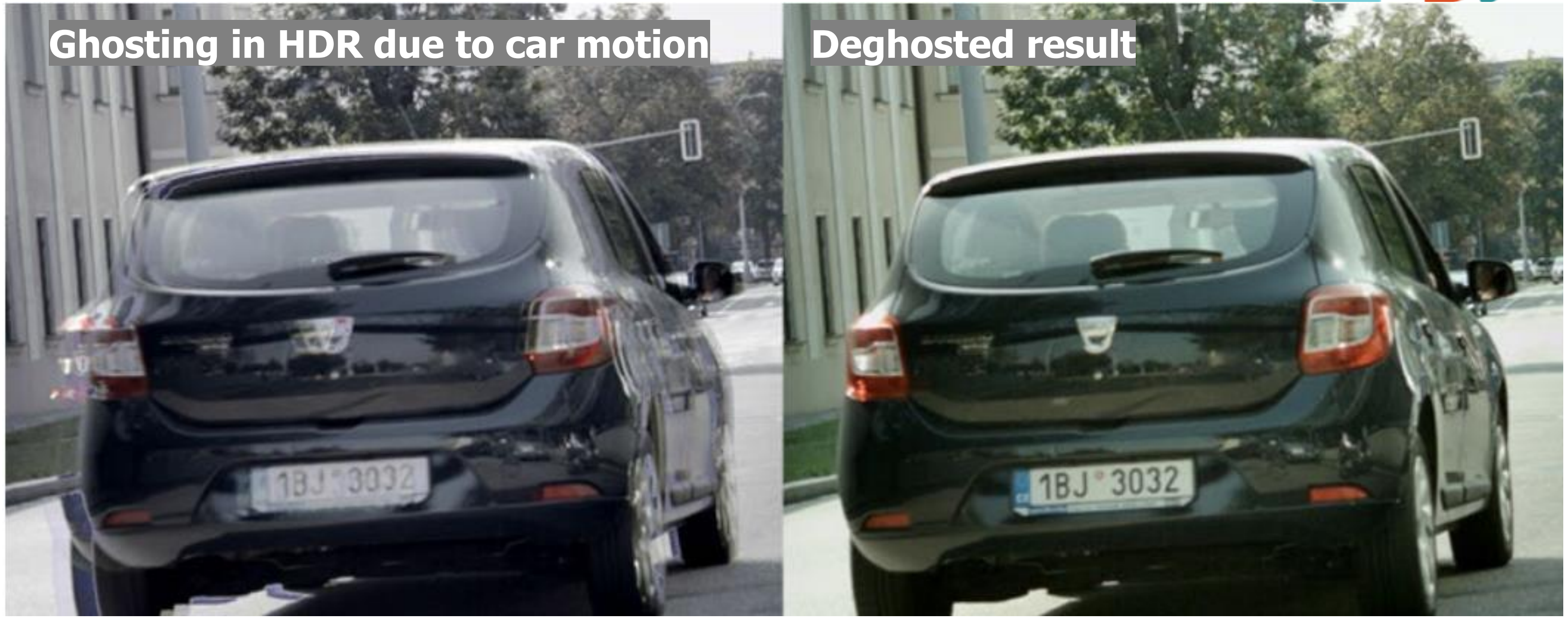




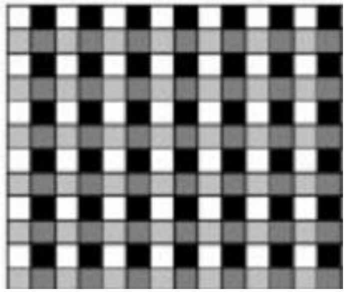
Images: Cecilia Aguerrebere et al.

HDR: deghosting and single-capture

- Multi-capture HDR can be deghosted
- Besides de-ghosting techniques, there are sensors that can produce HDR images from a single exposure. E.g.:
  - Differently-sized pixels
  - Spatially-varying filters



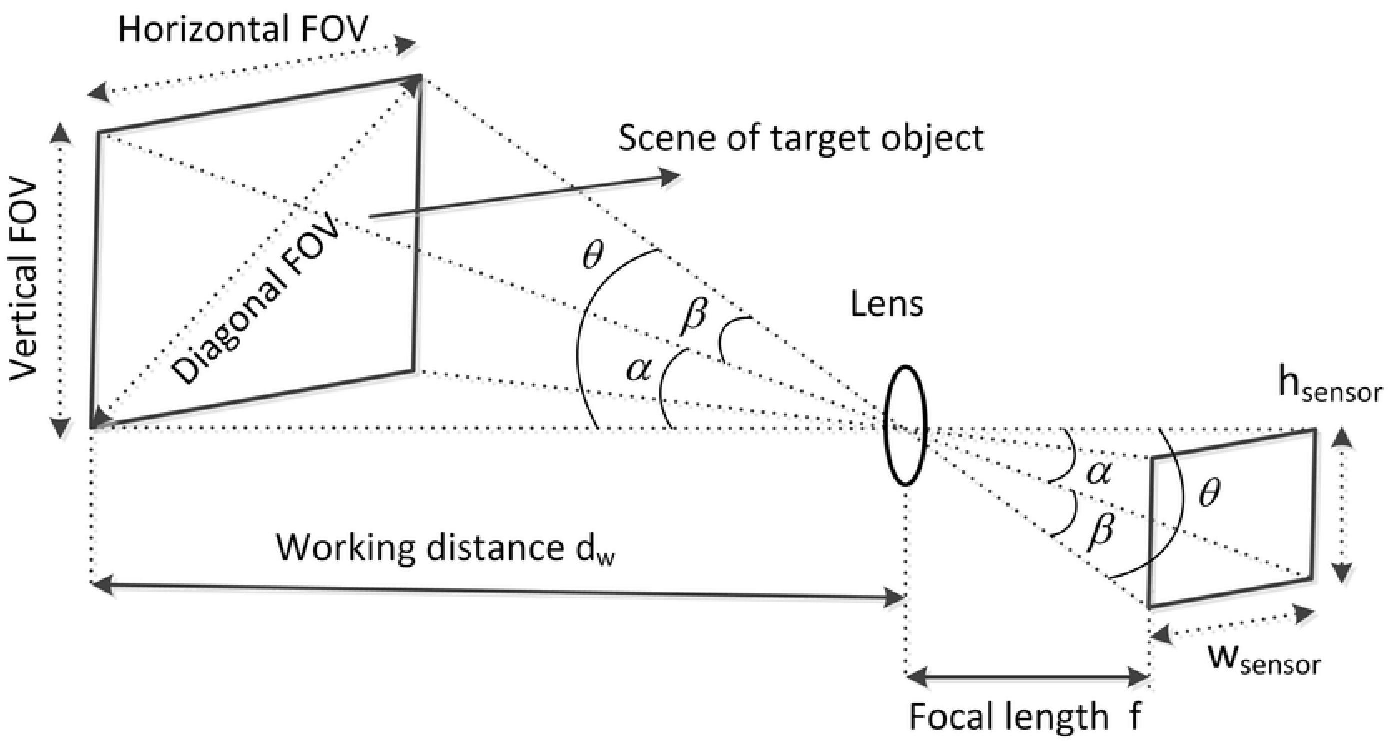
Source: Martin Musil et al.



Lenses

DigitalSense

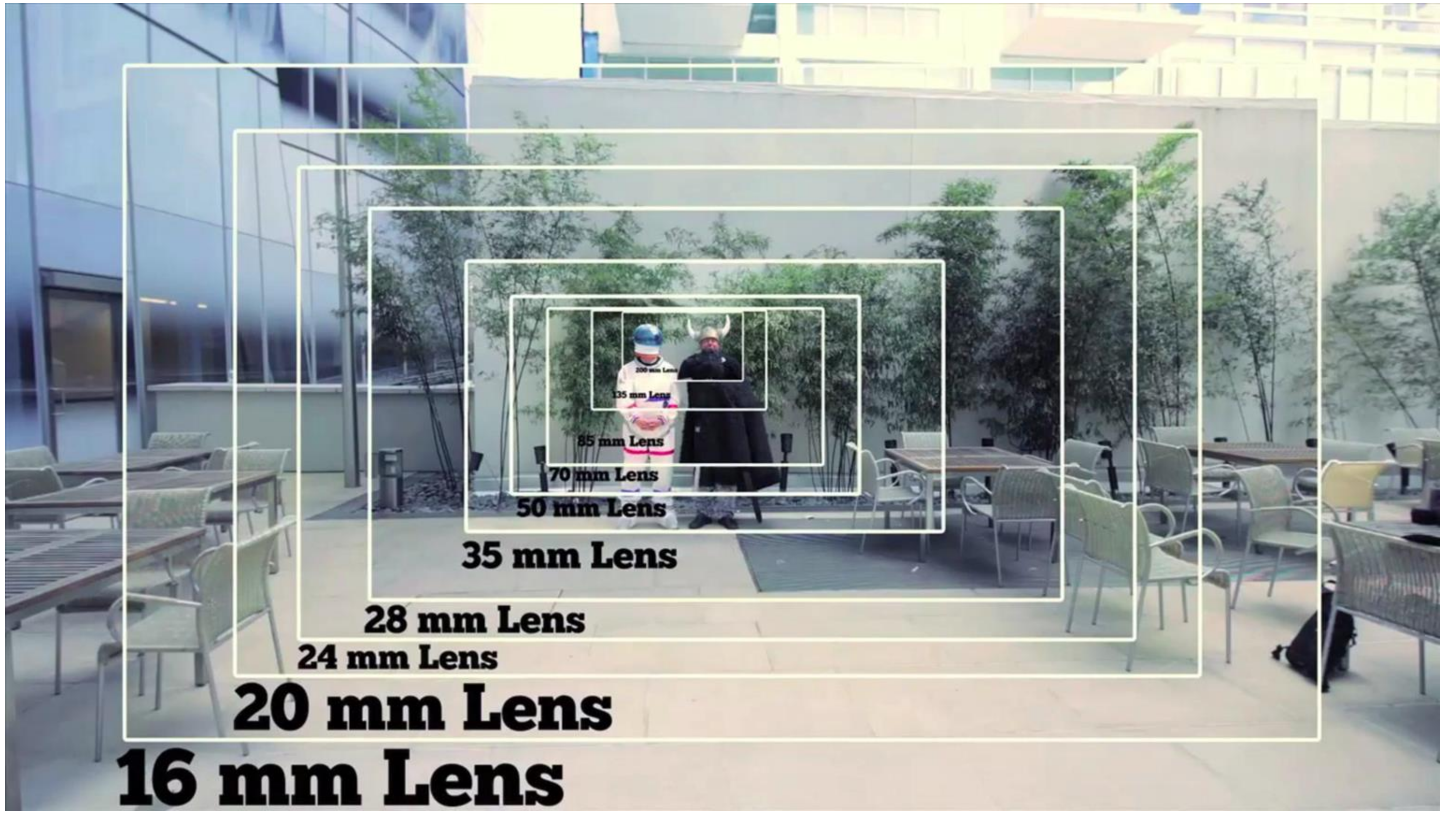
Lens basics



- **Field of view:** extent of the scene that is projected onto the sensor.
- **Focal length:** distance between the center of the lens and the sensor image plane.

Diagram via Janik Mabboux et. al

Focal length and field of view



- **Wider FOV:**
  - Lower resolution (1 pixel = large area)
- **Narrower FOV:**
  - Lower SNR (less photons per pixel)

Increase focal length: increase zoom, narrow FOV

Image: Vimeo Video School

Other considerations regarding lenses

- Fixed/variable focus
- Fixed/variable aperture
- Geometric distortion
- Chromatic aberration

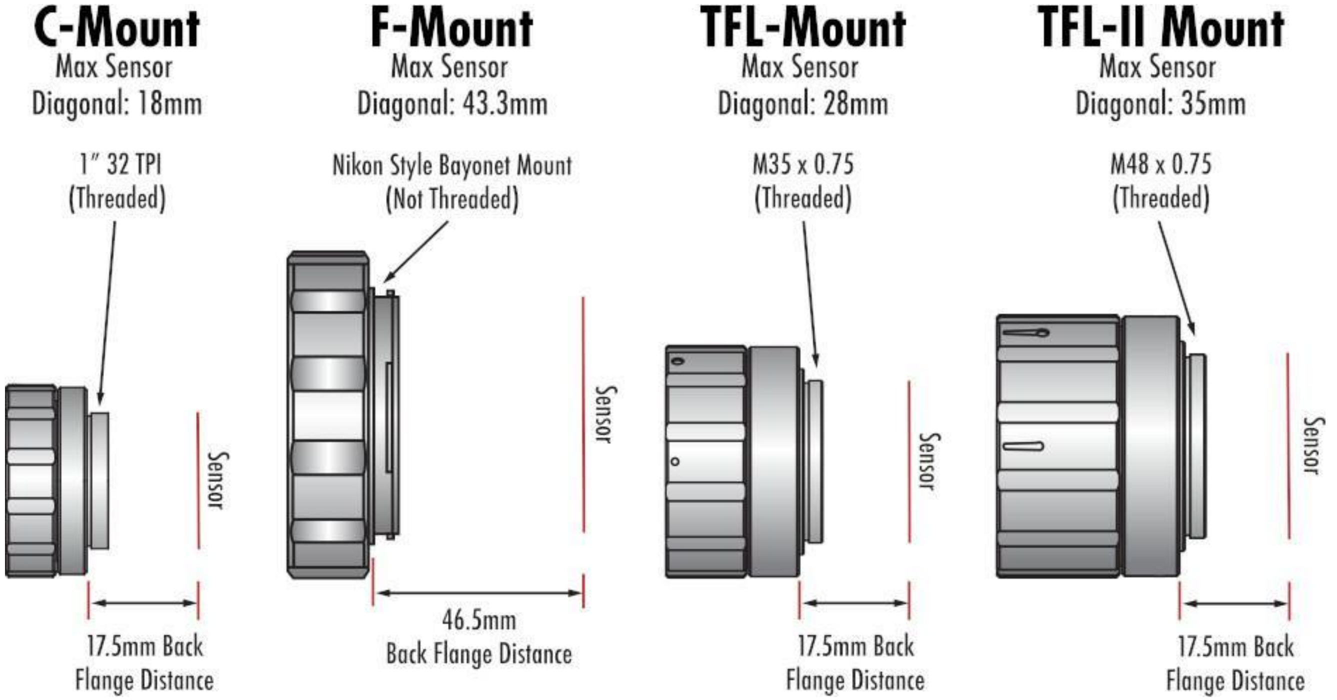
Images: Terpstra et al.

Images: Stan Zurek

Lens-sensor pairing

Images: Edmund Optics

- Sensors come in different sizes, and cameras provide different lens mounts



Lens-sensor pairing

- The lens' **resolution** should match the sensor's resolution.
- The lens' should match the **sensor size format**:
  - The image circle should cover the whole sensor.
  - Check the lens' specs to see suitable sensor sizes for the lens.
- The lens' and the sensor's **spectral properties** should match.

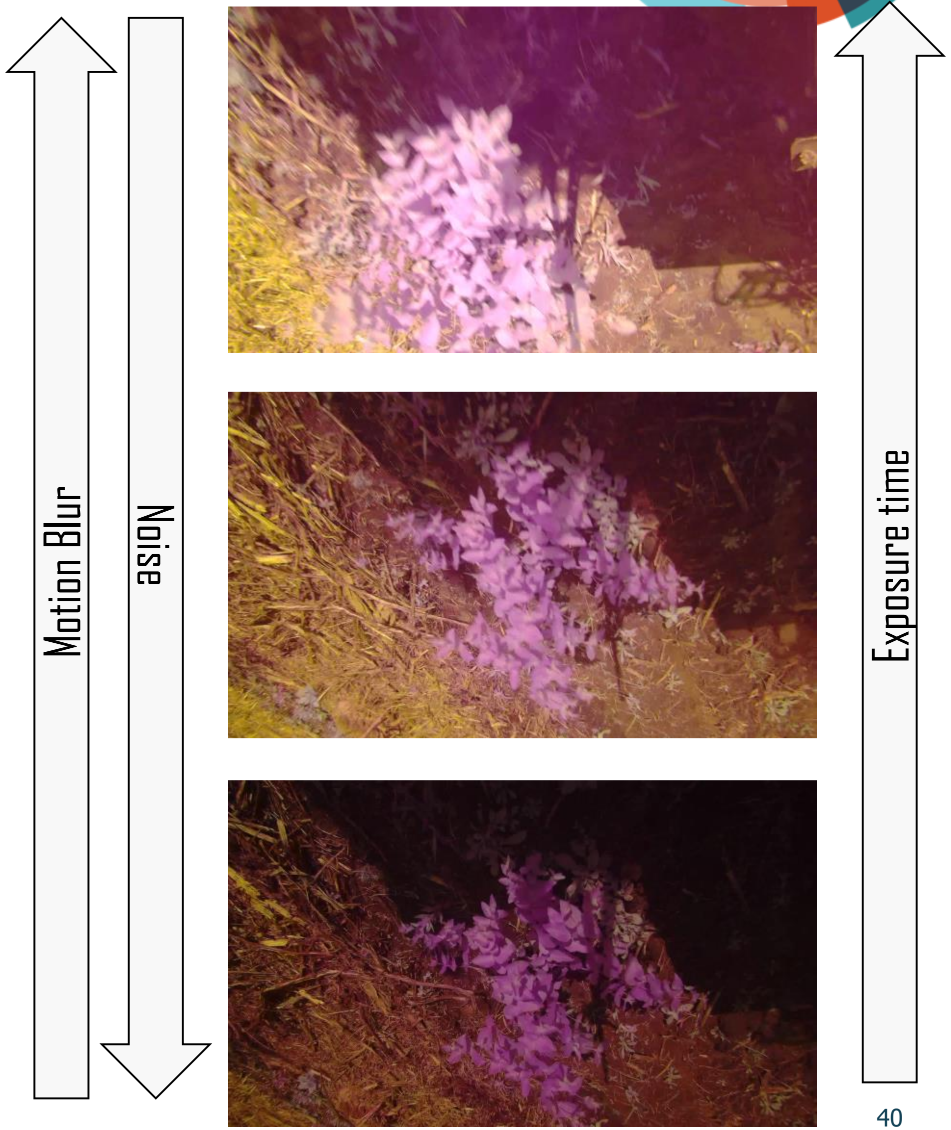
Illustration of image circle on sensor plane.

Image: nagualdesign

Configuration

Exposure time

- It's the length of time that the digital sensor inside the camera is exposed to light.
- Needs to be properly set for the application. Will have an effect on:
  - **Motion blur** (higher exposure -> more blur)
  - **Image noise** (lower exposure -> lower SNR)
  - **Under exposure and saturation**



Aperture, depth of field and exposure

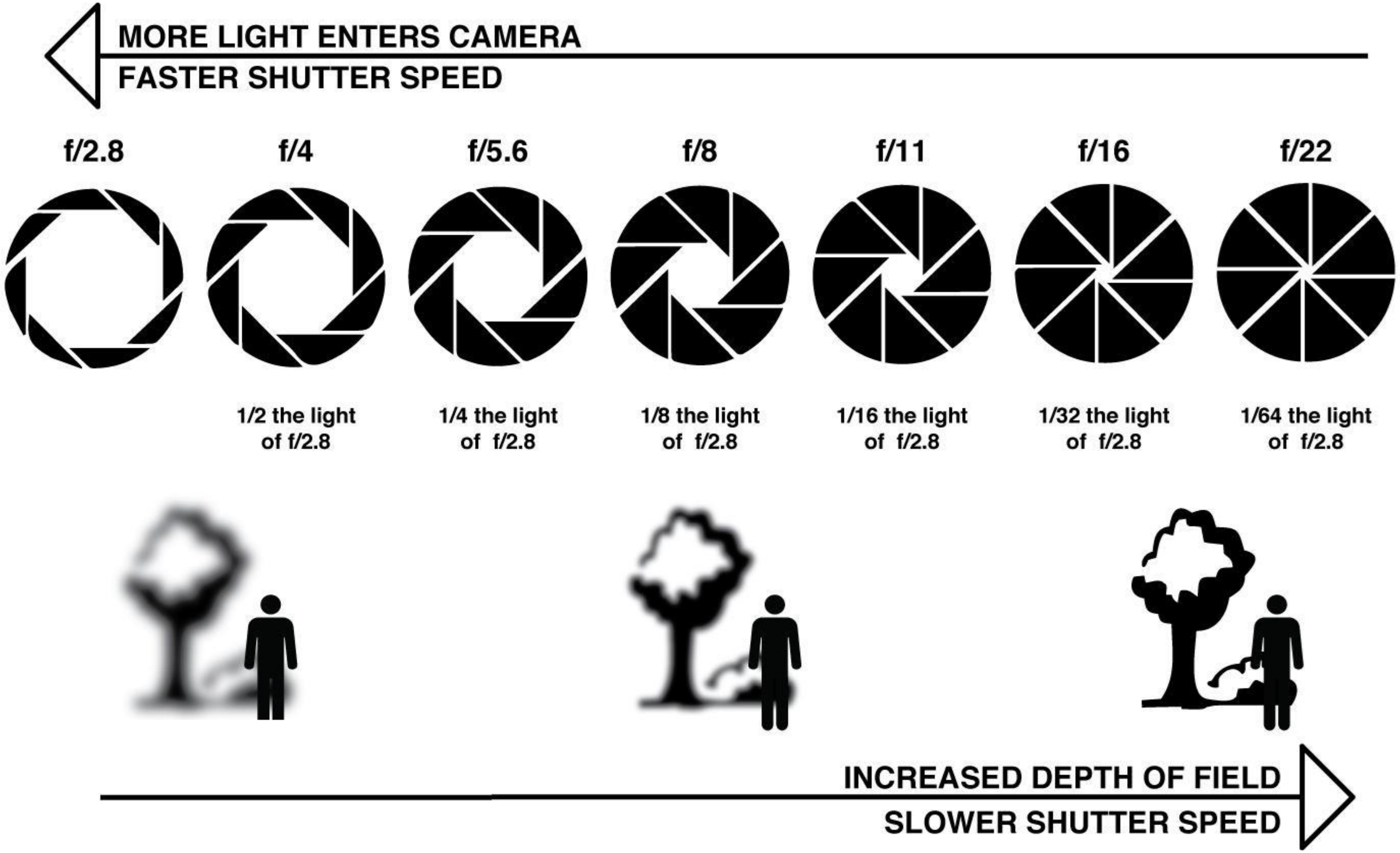


Diagram: Stuart Grais

White balance

- The image captured by the camera will mix colors from the scene with that of the light source.
- **White balance** post-processes the capture to try to compensate for the light source color and **reproduce the scene as observed under white light**.
- It can be difficult to know how algorithms operate in commercial cameras.

Images: Henry Maître

Software configurability

- Exposure, focus and white balance (to name some important ones) are **elements that you** will likely **need to adapt to your application**.

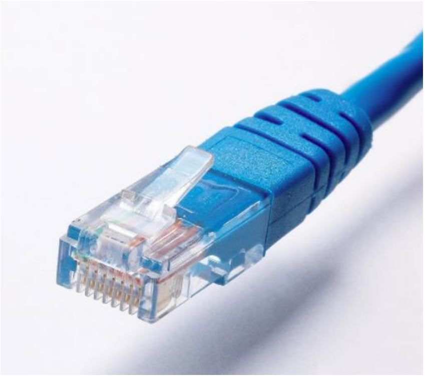
- **Specially during** the **prototyping** phase, the ability **configure** and tune them **via software** adds important flexibility.

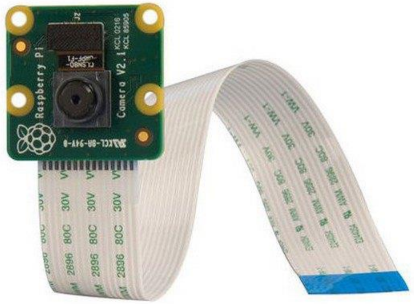
  - Which aspects of the camera can be controlled via software and how?
    - E.g. can you set a custom white balance preset?
    - or an auto white balance, exposure or focus on a ROI?

Interfaces & other considerations

DigitalSense

Interfaces





	USB 3.0 / 3.1	GigE	GMSL	MIPI CSI-2
Bandwidth	400 MB/s	125 MB/s	750 MB/s	1280 MB/s on four 320 MB/s data lanes
Max cable length	5 m (without active extenders)	100 m	15 m	30 cm
Main advantages	Highest ease of integration	Highest flexibility in camera placement	Good fit for automotive applications	Great bandwidth at smallest form factor

Other considerations

- Frame rate
- Compression
- Ruggedness
  - Dust & water rating
  - Shake & vibration
  - Temperature range
- External trigger
- Price
- Lead time
- Support
- QA

Conclusion and recommendations

DigitalSense

Conclusion and recommendations

- **If you want the best performance for your computer vision project, pay the proper attention to the acquisition system design.**
- Always try to control as many aspects of your capture as you can (lighting, distance to objects of interest, occlusions, etc.)
- List explicitly the camera features that are required for your project and which are desirable. Prioritize.

DigitalSense

Conclusion and recommendations

- Plan for camera integration time:
  - What's the interface: USB, MIPI?
  - Have you worked with this vendor's cameras before?
- Account for lead times in your project calendar
  - Prototype with the camera you have, but limit redoing work
- Establish good relationships with vendors you can rely on

Resources

Resource Slide

Maître, Henri. From photon to pixel: the digital camera handbook. John Wiley & Sons, 2017.

https://www.wiley.com/en-us/From+Photon+to+Pixel:+The+Digital+Camera+Handbook,+2nd+Edition-p-9781786301376

Edge AI + Vision Alliance cameras and sensors resources.

https://www.edge-ai-vision.com/resources/technologies/cameras-and-sensors/

Edmund Optics' knowledge center

https://www.edmundoptics.com/knowledge-center

Backup material

DigitalSense

Definitions

- **Field of view:** extent of the scene that is projected onto the sensor.
- **Working distance:** The distance from the front or first surface of the lens to the object under inspection.
- **Focal length:** distance between the center of the lens and the sensor image plane.
- **Depth of Field (DOF):** The maximum object depth that can be maintained entirely in acceptable focus.
- **F-number / f-stop:** the ratio of the system's focal length to the diameter of the entrance pupil.

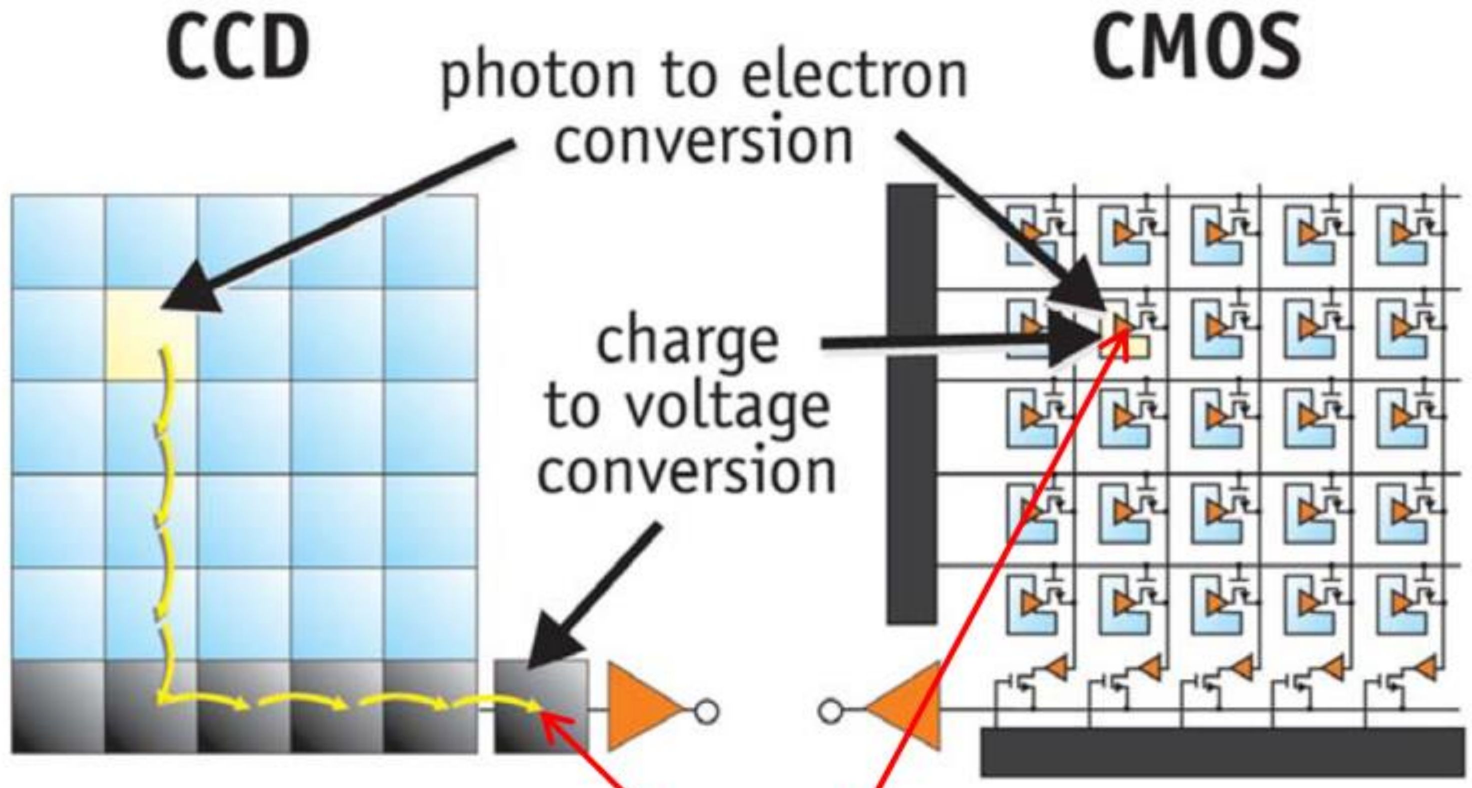
Definitions

- **Sensor resolution:** its dimension in pixels.
- **Camera resolution:** the minimum distance between two distinguishable points in an image, typically specified as a spatial frequency in units of line pairs per millimeter.
- **Quantum efficiency:** the measure of the effectiveness of an imaging device to convert incident photons into electrons.
- **Dynamic range:** the ratio between the largest non-saturating input signal and the smallest detectable input signal.

CCD vs. CMOS

- The 2 main sensor architecture types:
  - **CCD:** Charge-coupled device
  - **CMOS:** Complementary metal-oxide-semiconductor
- Historically, CCDs provided better image quality at higher cost
- **CMOS** with its **lower cost**, **lower energy usage** and **smaller form-factor** rode the smartphone wave and caught up with CCD in image quality
- CMOS is now dominating the embedded vision market, though CCDs may be a good fit for specific applications

CCD vs. CMOS



CCDs move photogenerated charge from pixel to pixel and convert it to voltage at an output node. CMOS imagers convert charge to voltage inside each pixel.

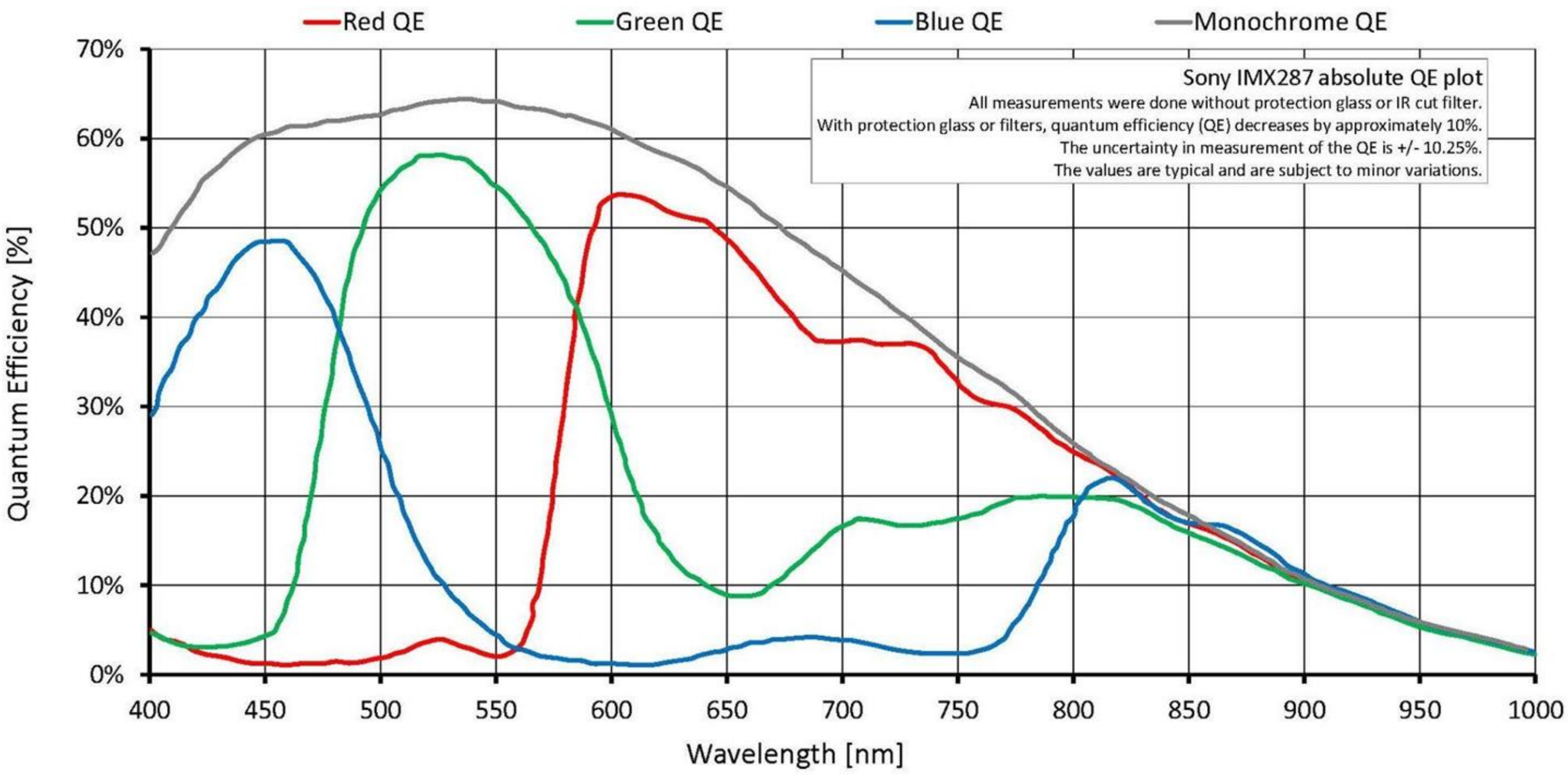
Read-out noise generated

Diagram via Gatan Inc.

Sensor quantum efficiency

- Monochrome has a higher efficiency than RGB

- Wavelengths over 710 nm are typically cut off with an IR-blocking filter



Aperture and depth of field

Image: Elizabeth Mott

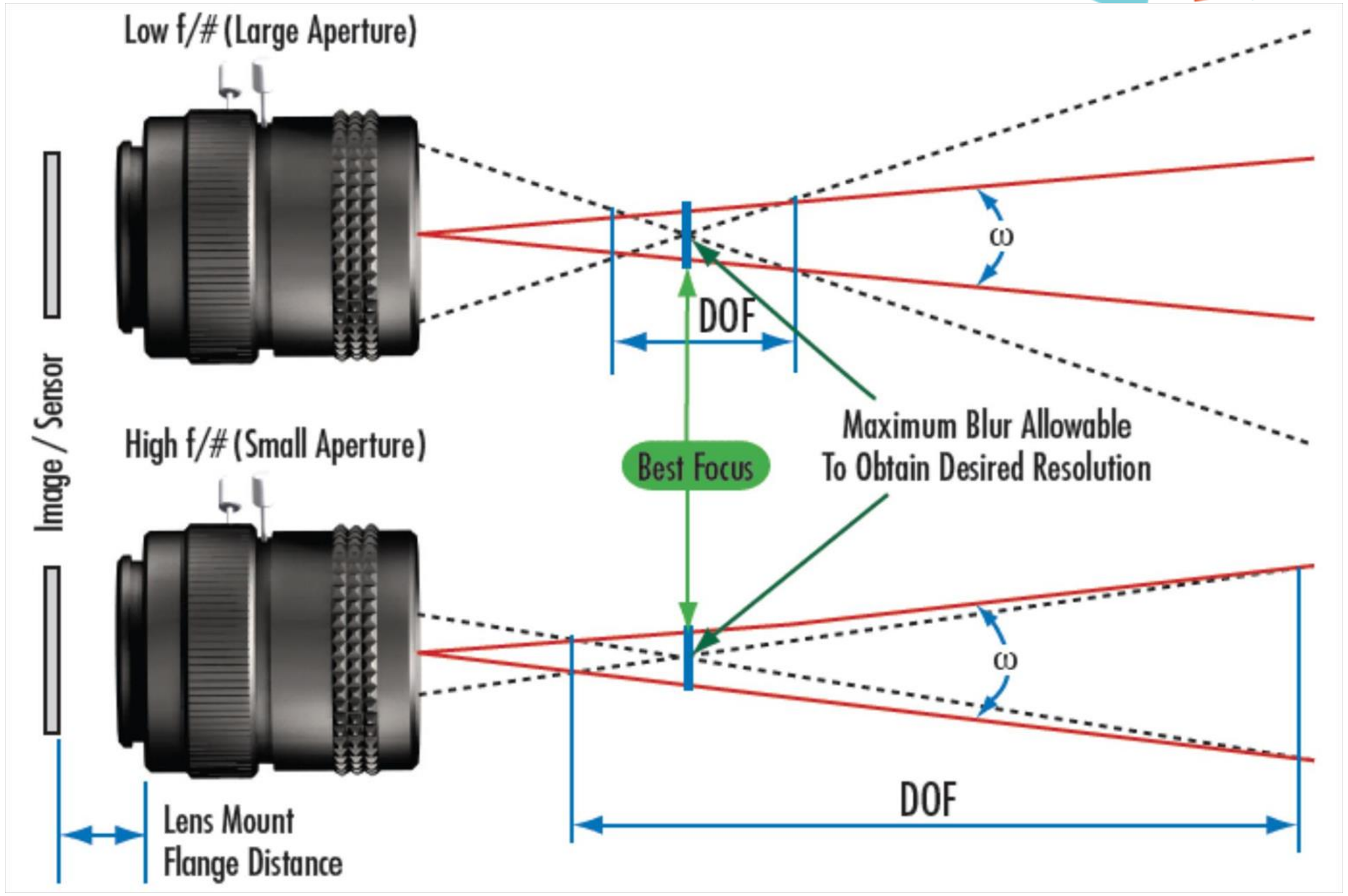


Diagram: Edmund Optics

Your Computer Vision and
Machine Learning R&D lab

Services

Low-risk POC

Develop a proof of concept using a process that marries the latest research with agile development.

Staff Augmentation

Our style of staff augmentation is based on teams that can self-manage to meet or exceed client expectations.

Component Development

We develop tailored state-of-the-art Computer Vision or Machine Learning components to integrate into your apps or systems.

Whole Application Delivery

We build full web systems or apps for customers that prefer all development in the project to be handled by a single trusted vendor.

DigitalSense

Let's discuss your next project!

We'd love to hear from you
hello@digitalsense.ai

digitalsense.ai