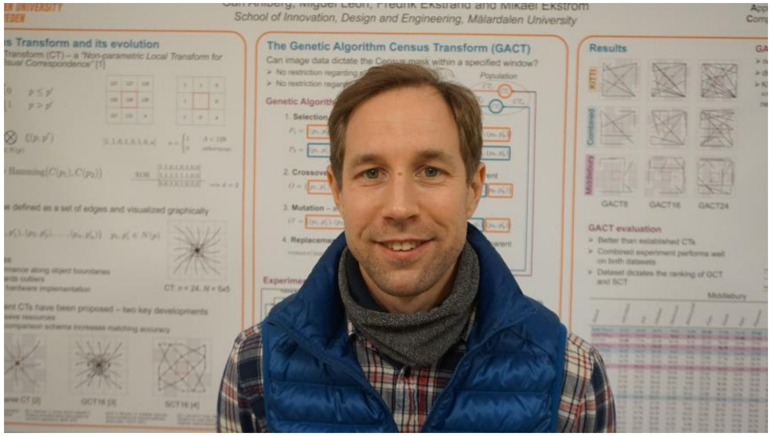
Camera System

DVA138 2023

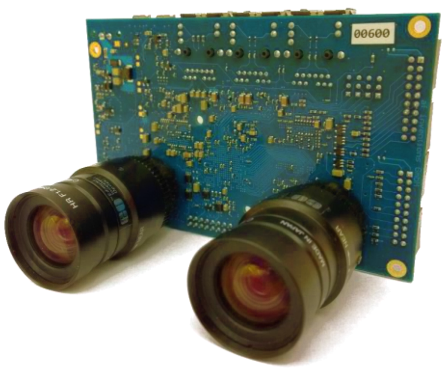


Carl Ahlberg carl.ahlberg@mdh.se

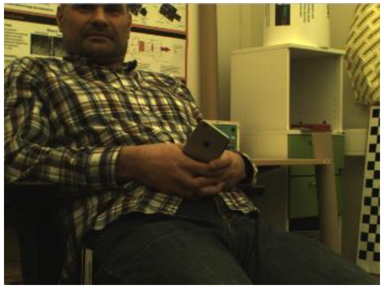
MÄLARDALEN UNIVERSITY SWEDEN

Our Research in Image Processing/Computer Vision

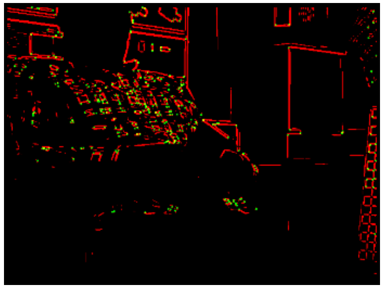
- **Carl Ahlberg** & Fredrik Ekstrand
- Thesis: Embedded high-resolution stereo-vision of high frame-rate and low latency through FPGA-acceleration
- Image algorithms in embedded systems and reprogrammable hardware (FPGA)
- Optimize speed and throughput
- Stereo camera system (GIMME)
- Genetic Algorithm for matching process optimization



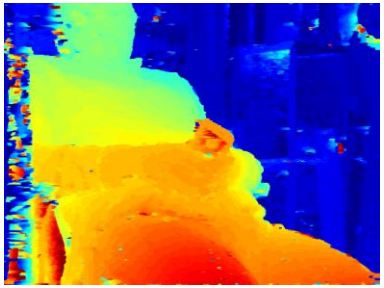
GIMME2



Image(s)



Harris



Stereo

(Digital) Camera System

Lens
- Distance to scene
- Amount of light
- Distortion

Sensor
- Resolution
- Data rate
- Sensitivity

World scene

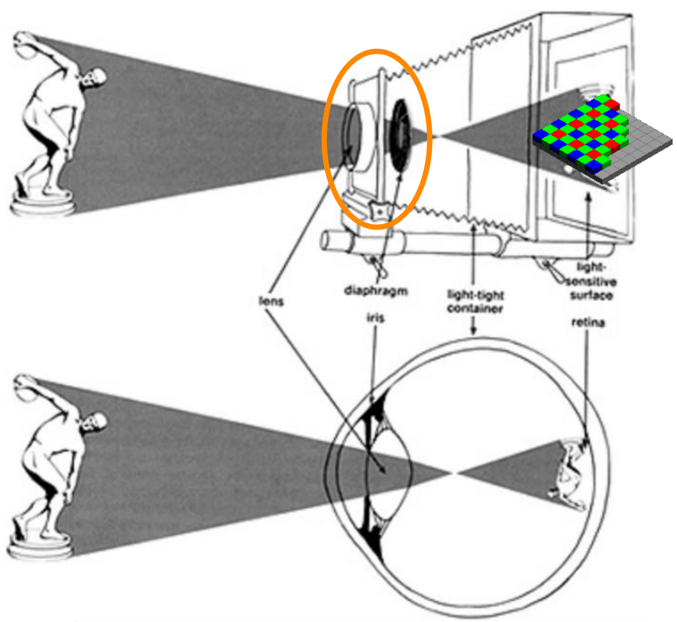
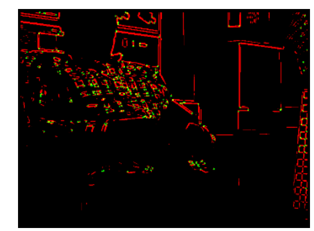
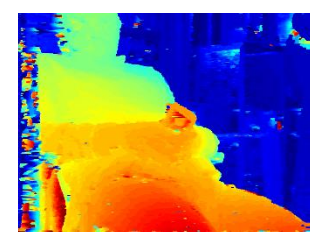


Image Processing
- Noise filtering
- Undistortion
- Color correction

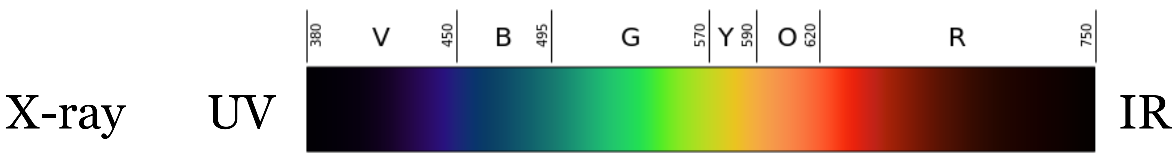
Computer Vision
- Extract application specific information from image
- Objects, features, depth, etc.



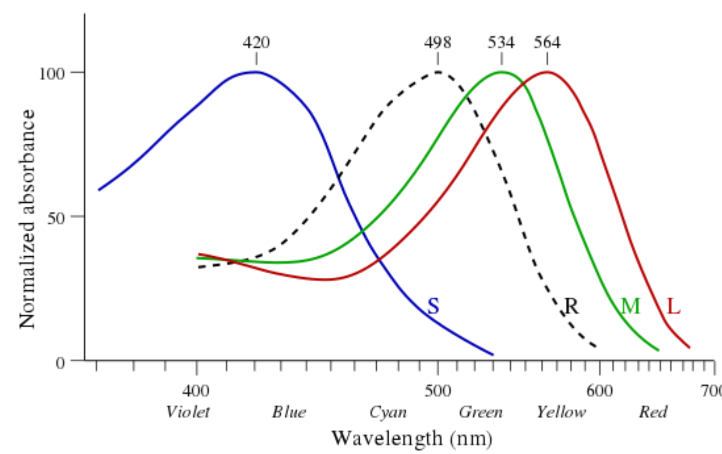


Visible light

- Light is electromagnetic radiation – light rays
  - Direct – light source = illumination
  - Reflection, refraction, (diffraction, dispersion)
- Visible spectrum: 380-750 nm



- IR and near-IR, thermal and night vision
- UV and X-ray
- Humans are trichromats
  - Some are dichromats
  - Some are tetrachromats
  - Perception differs



Images from Wikipedia

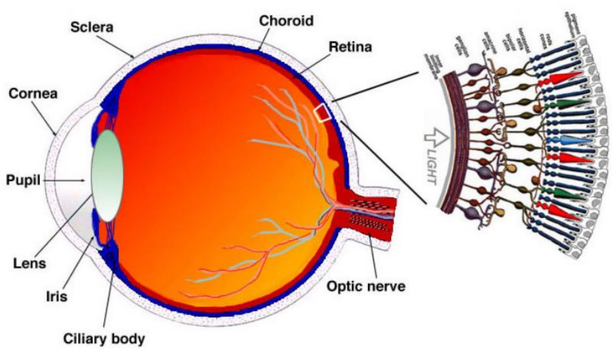
Vision

- What can you detect with vision?
  - Light
  - Shadows
  - Colors
  - Shapes
  - Objects
  - Patterns
  - Motion
  - Depth
- What does motion require?
  - Subsequent images?
- What does depth require?
  - Different views?
  - Two cameras?

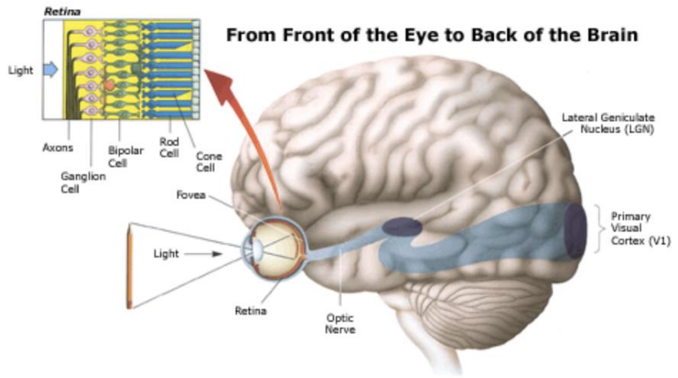
One picture, a thousand words

- **Vision** is our most powerful sense in aiding our perception of the 3D world around us.
- Retina is ~10cm^2. Contains millions of **photoreceptors**
  (120 mil. rods and 7 mil. Cones for colour sampling)
- Provides **enormous** amount of information: data-rate of **~3 *GBytes/s***

⇨ a large proportion of our brain power is dedicated to processing the signals from our eyes

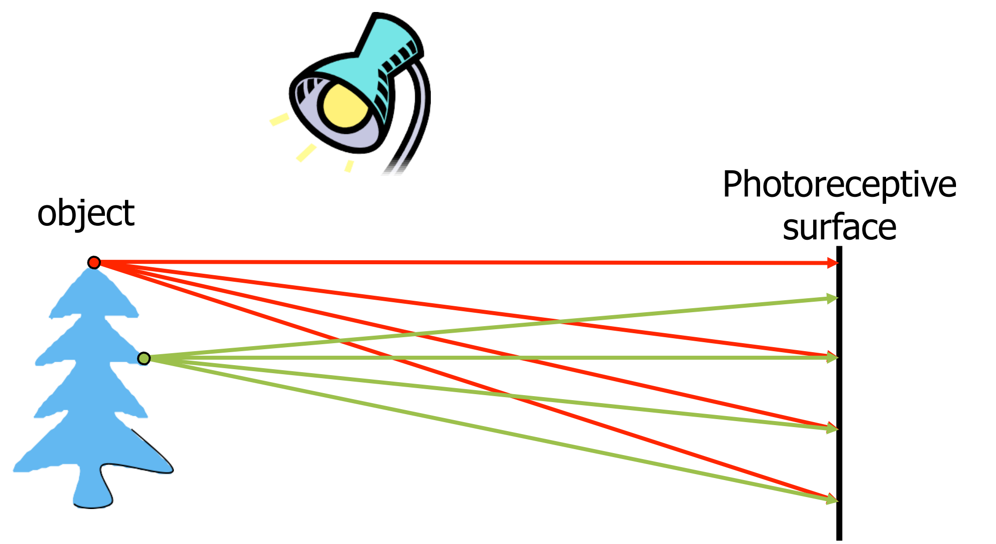


http://webvision.med.utah.edu/sretina.html



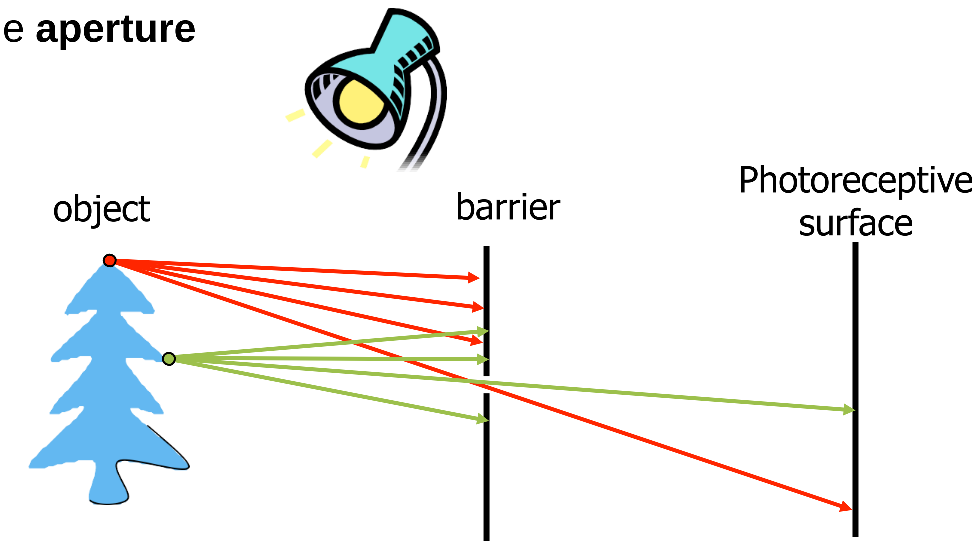
The camera | image formation

- If we place a piece of film in front of an object, do we get a reasonable image?



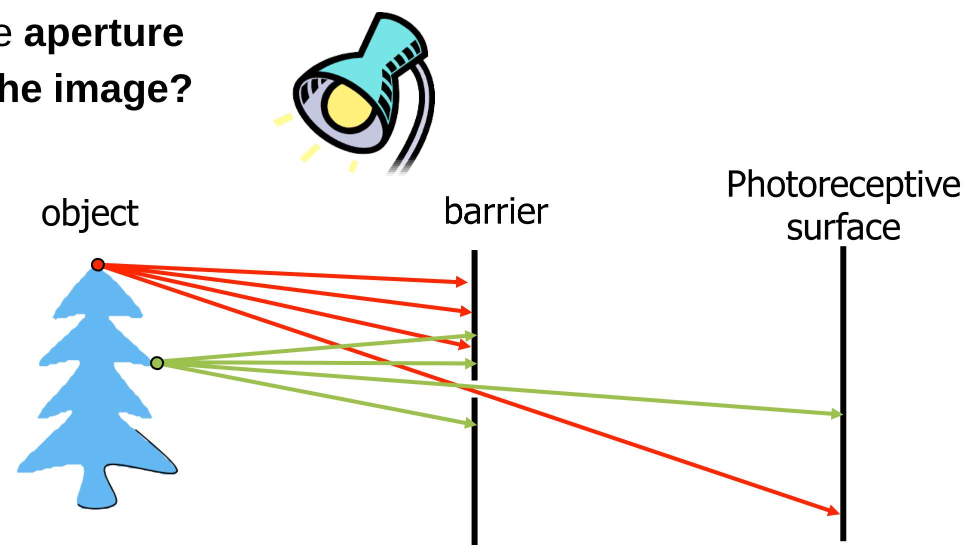
The camera | image formation

- If we place a piece of film in front of an object, do we get a reasonable image?
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**

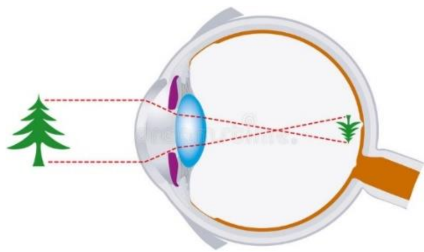


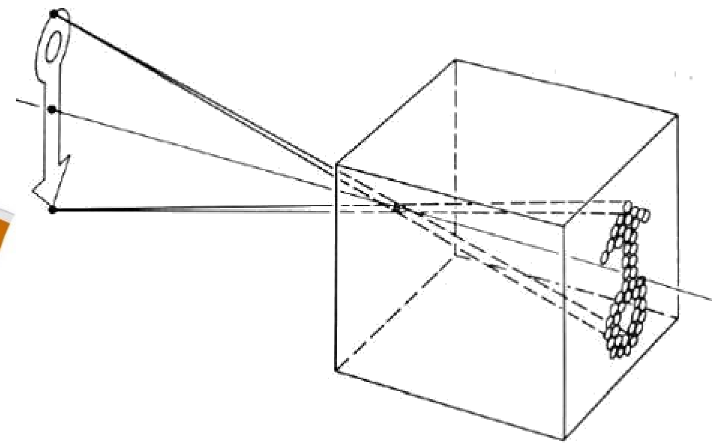
The camera | image formation

- If we place a piece of film in front of an object, do we get a reasonable image?
- Add a barrier to block off most of the rays
  - This reduces blurring
  - The opening is known as the **aperture**
  - **How does this transform the image?**



The camera | the pinhole camera model





- Pinhole model:
  - Captures **beam of rays** – all rays through a single point (note: no lens!)
  - The point is called **Center of Projection** or **Optical Center**
  - The image is formed on the **Image Plane**

- We will use the pinhole camera model to describe how the image is formed

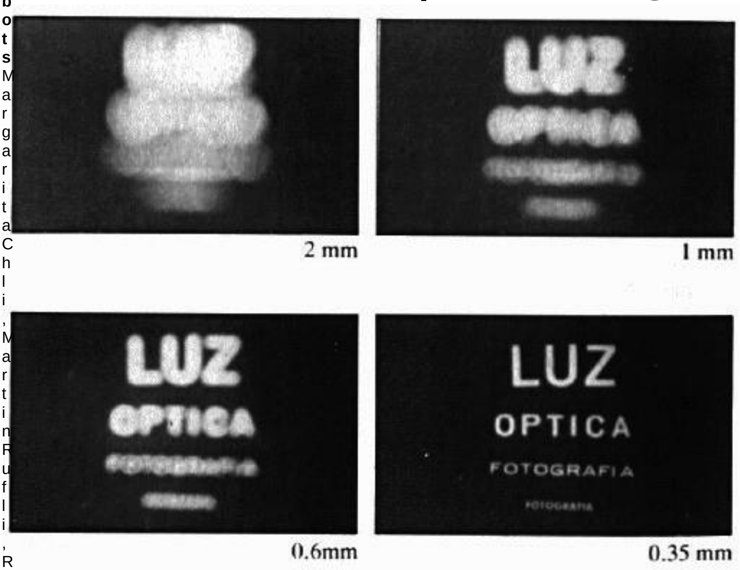
Based on slide by Steve Seitz

The camera | home-made pinhole camera

What can we do
to reduce the blur?

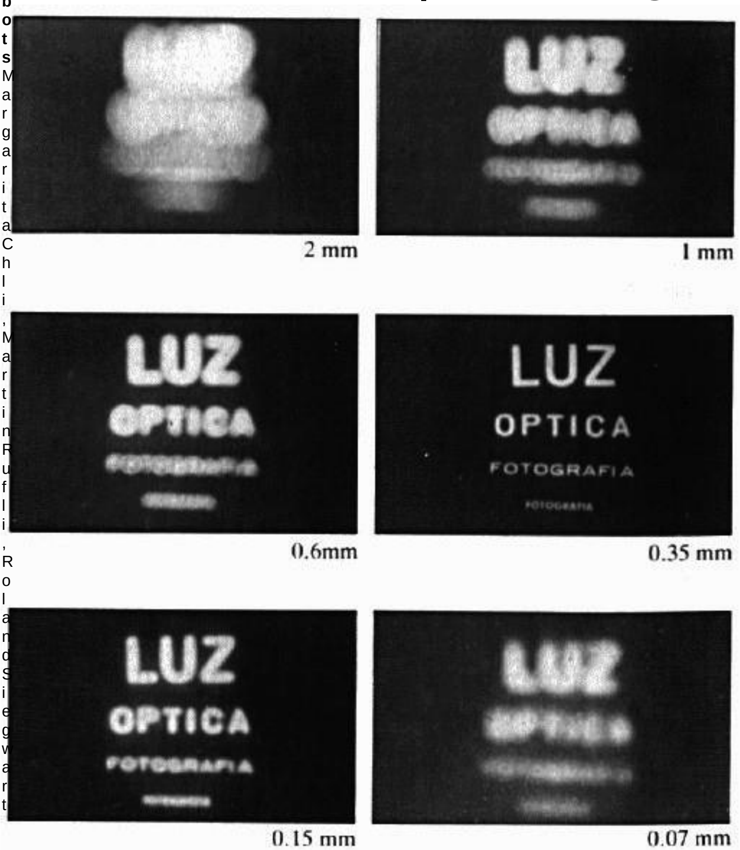
Based on slide by Steve Seitz

The camera | shrinking the aperture



Why not make the aperture as small as possible?

The camera | shrinking the aperture



Why not make the aperture as small as possible?

- Less light gets through (must increase the exposure)
- Diffraction effects…

AutonomousMobileRobots Margarita Chli, Martin Rufli, Roland Siegwart

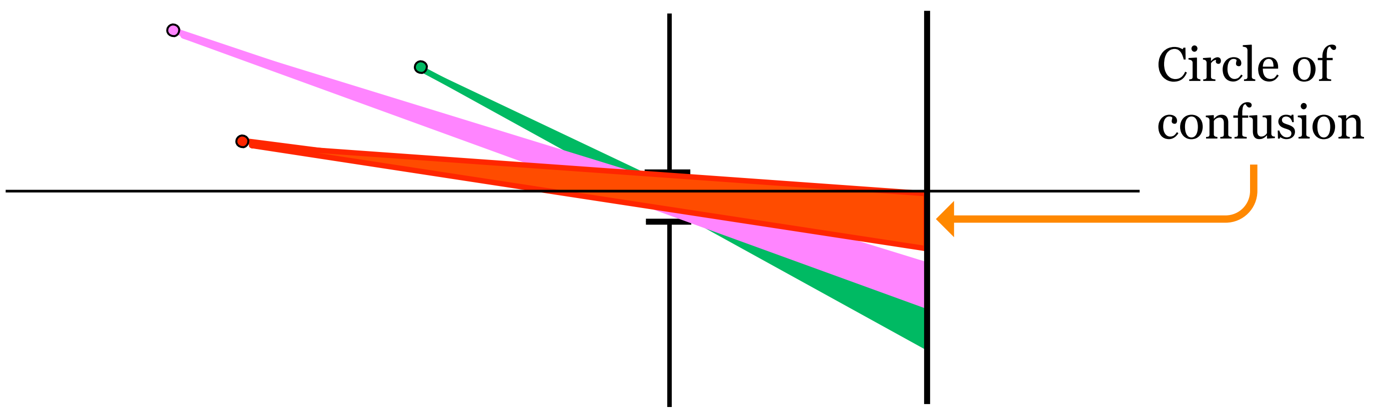
MÄLARDALEN UNIVERSITY SWEDEN

The camera | why use a lens?

VISION FOR ROBOTICS LAB

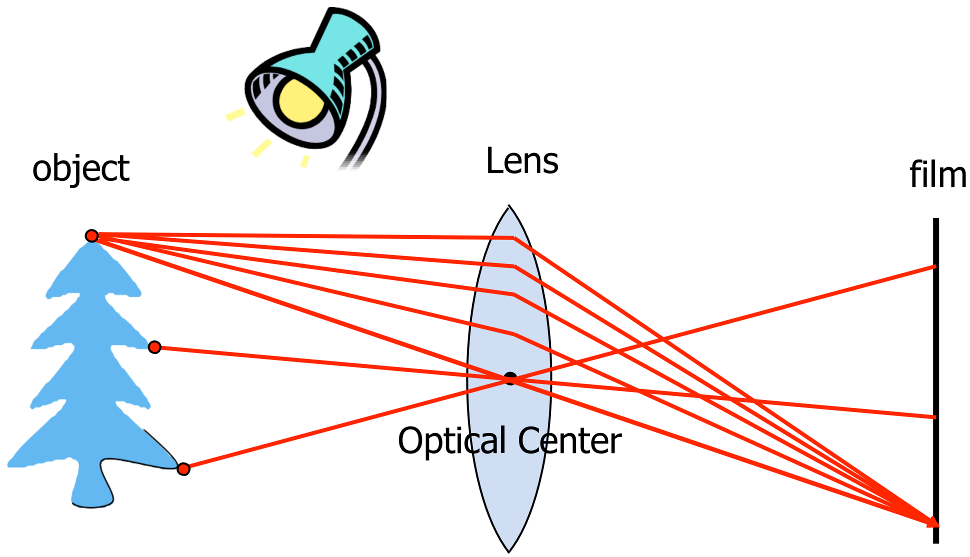
- *The ideal pinhole*:
  only one ray of light reaches each point on the film
  $\Rightarrow$ image can be very dim; gives rise to diffraction effects

- Making the pinhole bigger (i.e. aperture) makes the image blurry



The camera | why use a lens?

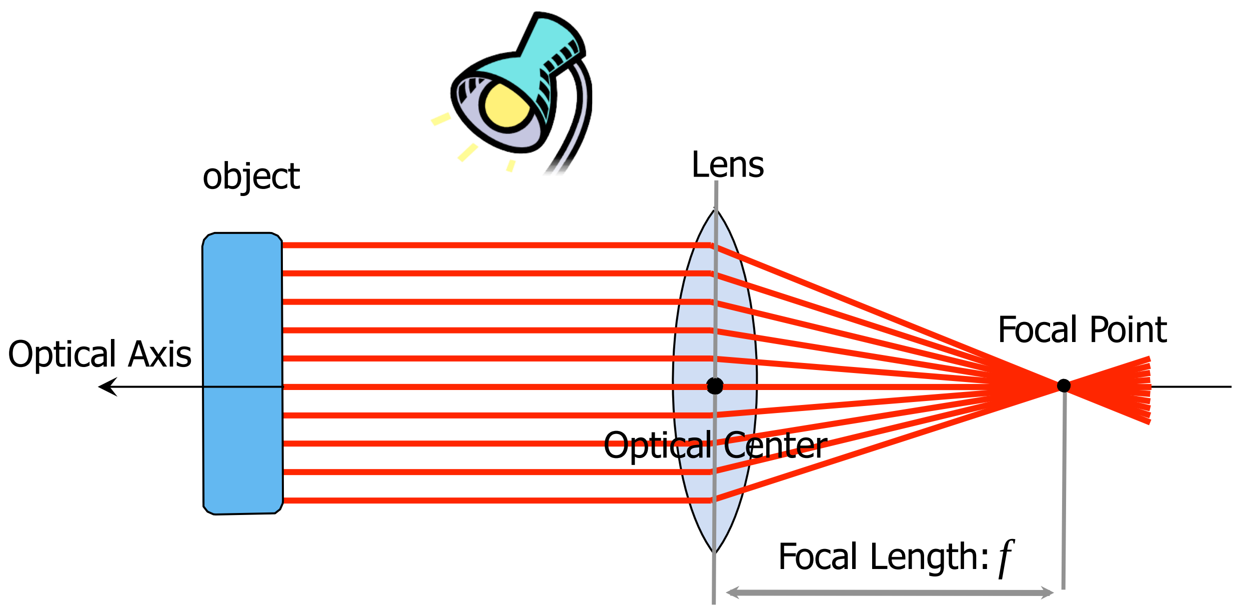
- A lens focuses light onto the film
- Rays passing through the **optical center** are not deviated



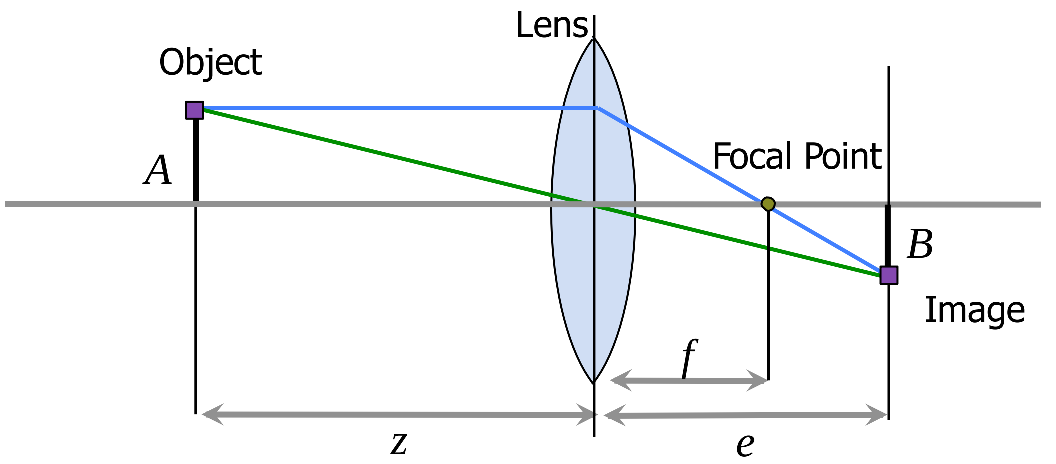
Autonomous Mobile Robots Margarita Chli, Martin Rufli, Roland Siegwart

The camera | why use a lens?

- A lens focuses light onto the film
- Rays passing through the **optical center** are not deviated
- All rays parallel to the **optical axis** converge at the **focal point**

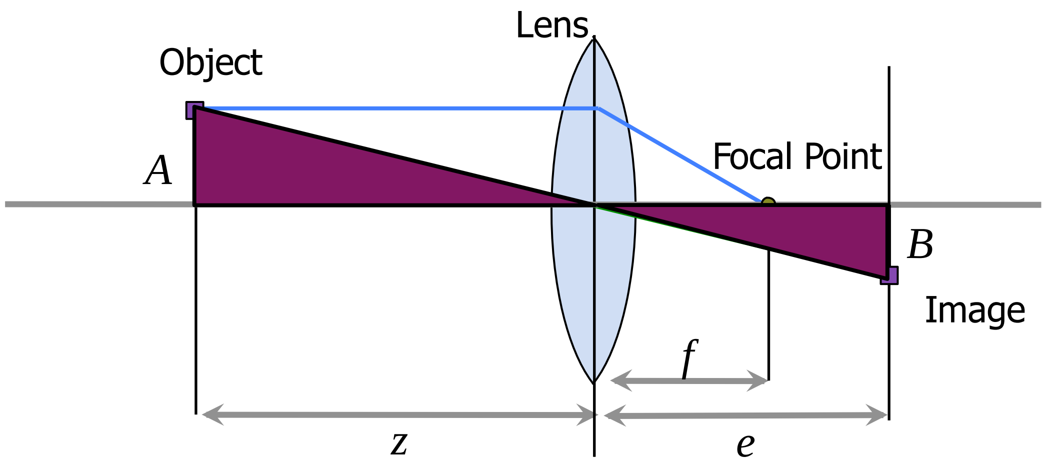


The camera | how to create a focused image?



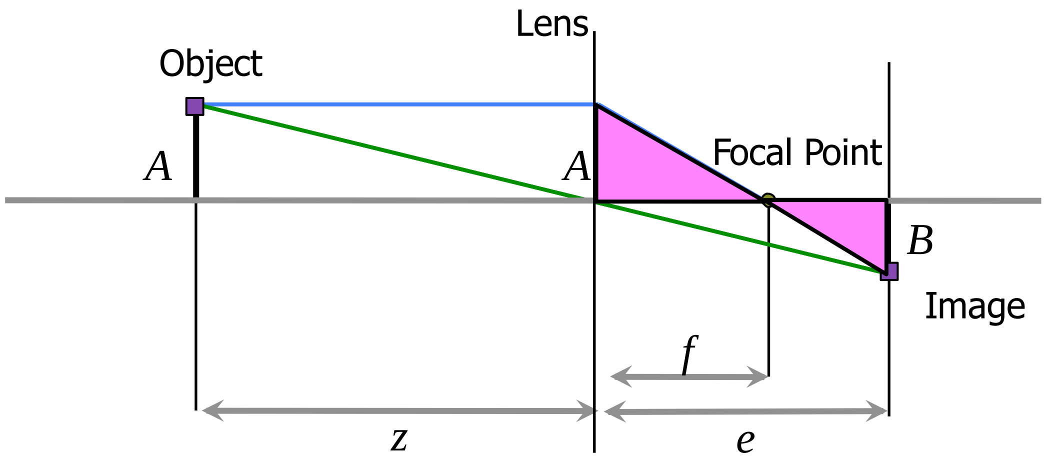
Find a relationship between f, z and e

The camera | the thin lens equation



- Similar Triangles: $\frac{B}{A} = \frac{e}{z}$

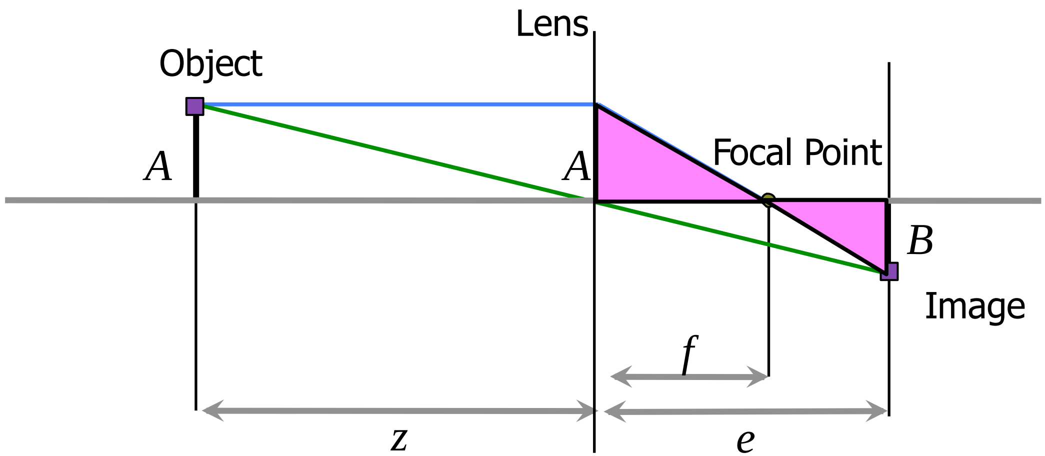
The camera | the thin lens equation



- Similar Triangles: $\dfrac{B}{A} = \dfrac{e}{z}$

  $\dfrac{B}{A} = ?$

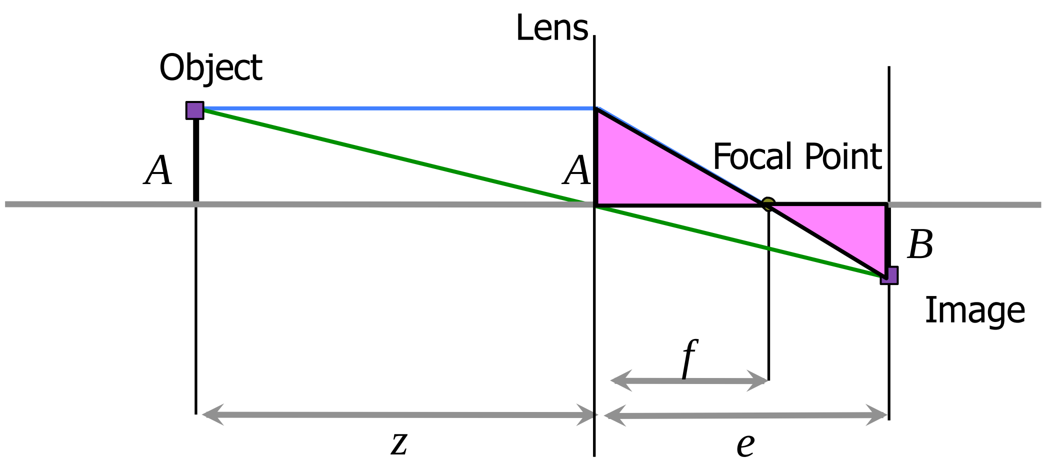
The camera | the thin lens equation



- Similar Triangles: $\frac{B}{A} = \frac{e}{z}$

$$\frac{B}{A} = \frac{e-f}{f} = \frac{e}{f} - 1$$

The camera | the thin lens equation

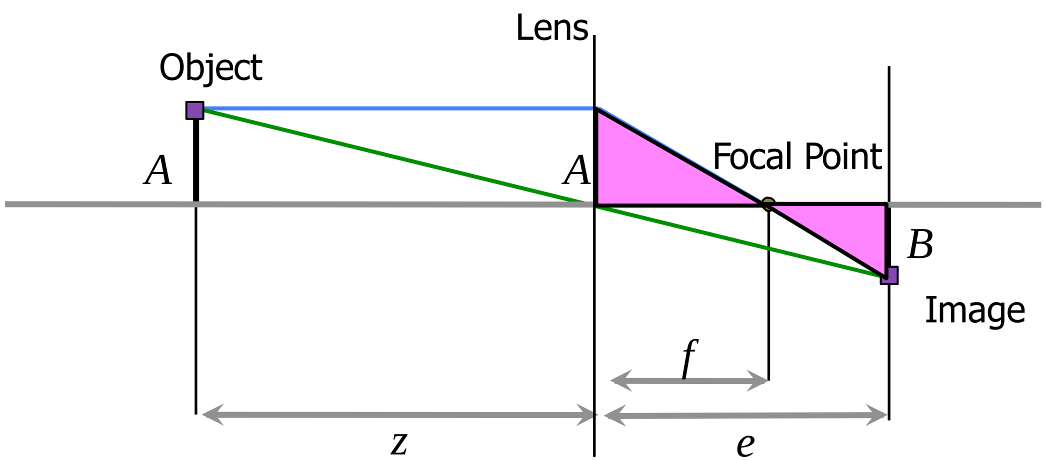


- Similar Triangles:

$$\left.\begin{aligned} \frac{B}{A} &= \frac{e}{z} \\ \frac{B}{A} &= \frac{e-f}{f} = \frac{e}{f} - 1 \end{aligned}\right\} \frac{e}{f} - 1 = \frac{e}{z} \Rightarrow \boxed{\frac{1}{f} = \frac{1}{z} + \frac{1}{e}}$$

"Thin lens equation"

The camera | the thin lens equation



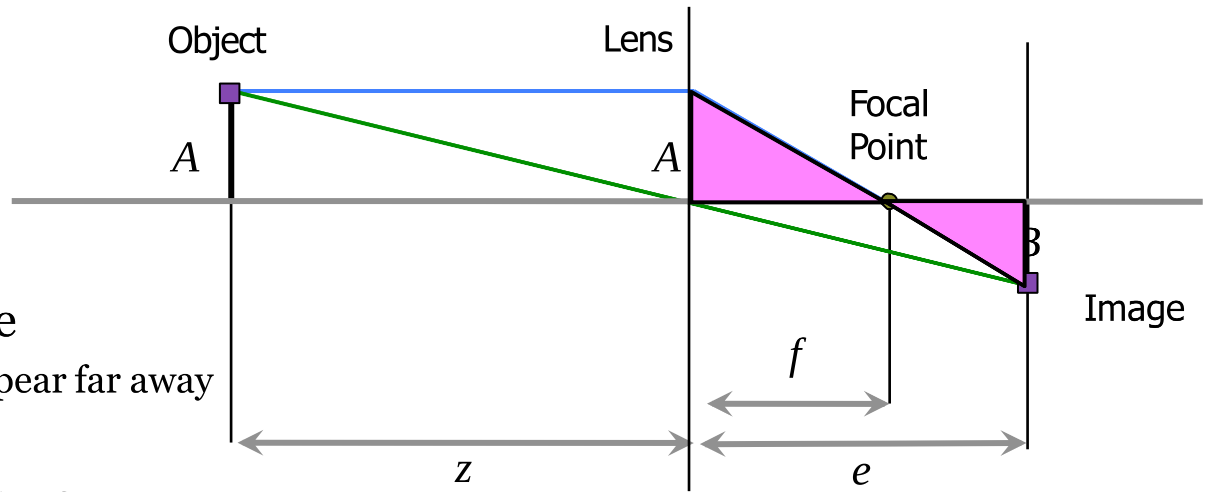
- Similar Triangles:

$$\left.\begin{aligned}\frac{B}{A} &= \frac{e}{z}\\ \frac{B}{A} &= \frac{e-f}{f} = \frac{e}{f} - 1\end{aligned}\right\} \frac{e}{f} - 1 = \frac{e}{z} \Rightarrow \boxed{\frac{1}{f} = \frac{1}{z} + \frac{1}{e}}$$

"Thin lens equation"

- Thus, when an object is in focus, the distance (range) can be calculated: Depth from focus.

The camera | lens properties



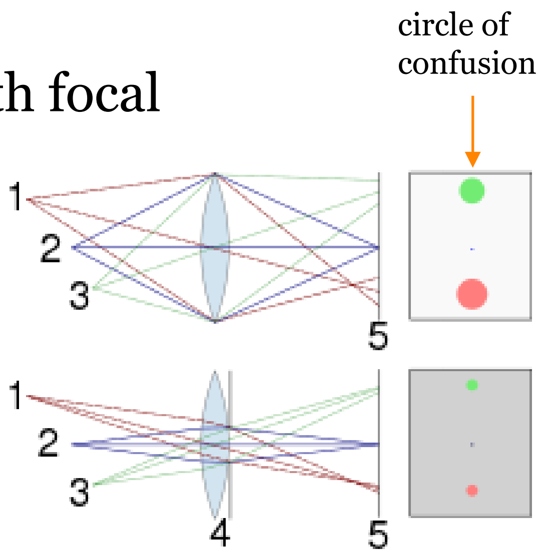
- ***f*** – focal length
  - Field of view (FOV)
  - Small ***f*** – wide angle
    - large FOV, objects appear far away
  - Large ***f*** – tele lens
  - Zoom lens – variable ***f***

Image from Wikipedia

- For embedded systems
  - Fixed lens
  - Determine FOV depending on application
  - Digital zoom = image processing

The camera | lens properties

- Aperture – *f-number* (not to be confused with focal length), f/1.8 or F=1.8
  - Pupil
  - Depth of field (DOF)
    - Threshold circle of confusion
  - Large aperture (small F)
    - Allows more light through => brighter image
    - Shallow depth of field



Images from Wikipedia

- For emedded systems
  - Fixed aperture – maximize DOF

Aperture = f/1.4. DOF=0.8 cm | Aperture = f/4.0. DOF=2.2 cm | Aperture = f/22. DOF=12.4 cm

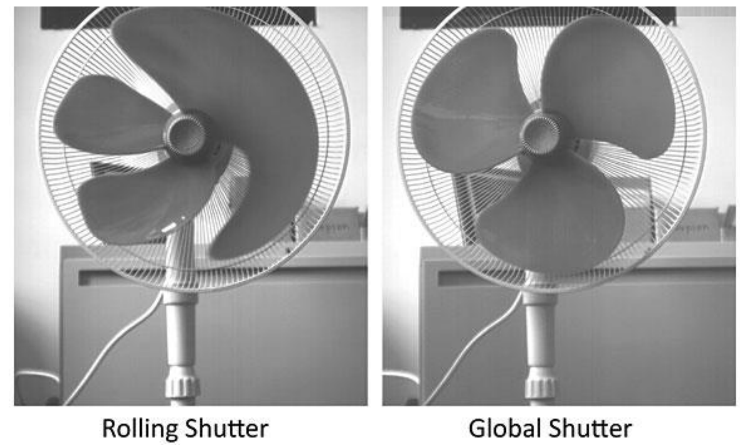
Shutter

- The shutter controls the time the sensor is exposed to light
  - Shutter speed
- Mechanical or electronic (mobile devices)
- An image is an integration of the light captured by the sensor over the time the shutter is open
  - Fast shutter speed => less light
    - Underexposed, dark images
  - Slow shutter speed => more light
    - Overexposed, bright images
    - Motion blur
    - Noise for long exposures in low light
- Relationship with Aperture in photography

Image from Wikipedia

Hardware considerations

- Electronic Rolling Shutter (ERS)
  - Common for low-cost CMOS
  - Image skew for fast moving object or ego-motion
  - Should be considered when developing computer vision algorithms for mobile devices



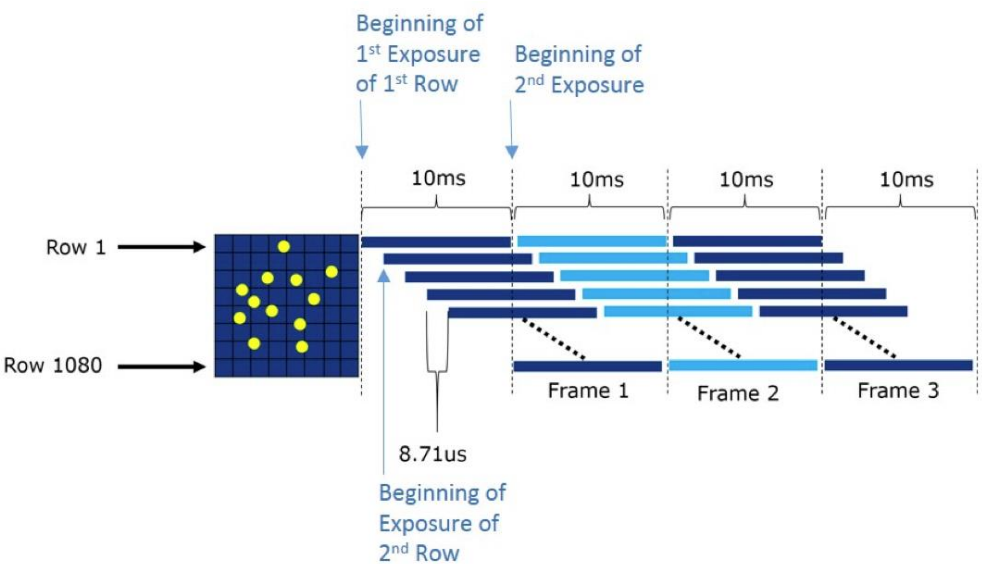
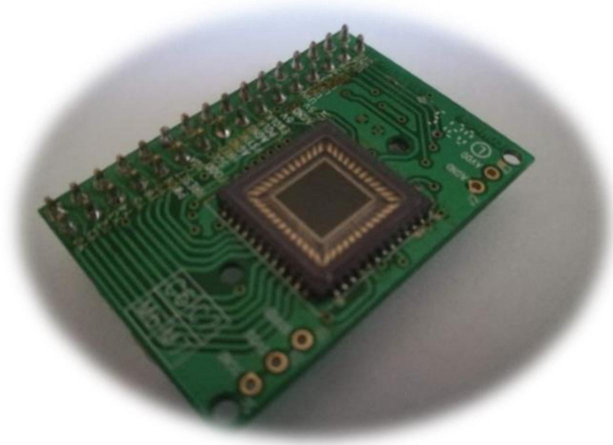


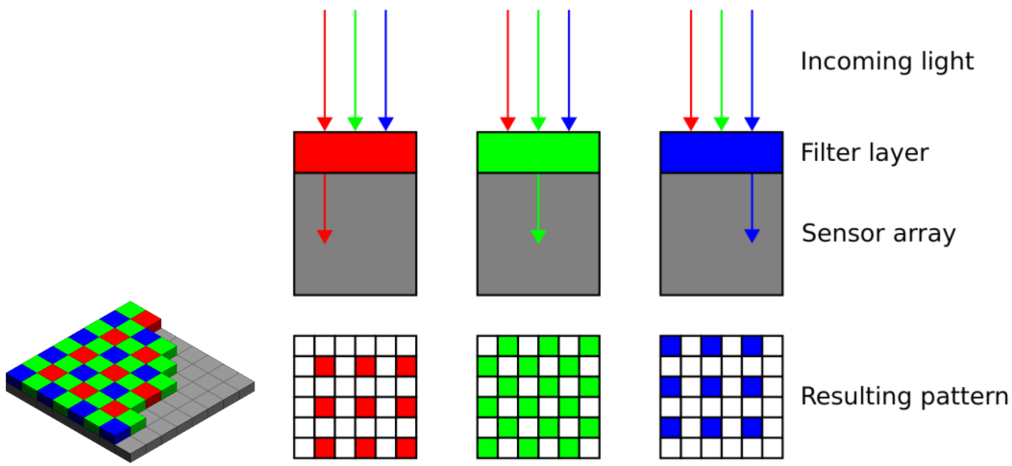
Image sensor



- 2D-array of photo diods
  - each diod represents one pixel
  - Pixel = picture element
- Image sensor properties
  - Resolution (confusing) = number of pixels (**spatial**)
  - Color resolution (digital ouput) = number of bits (**information**)
  - Output frequency (**temporal**)
    - Output resolution vs frames per second (fps)
  - Size => Pixel size
    - Larger sensors/pixels => more light => less noise
  - SNR, cost, power efficiency, built in functionality, etc.
- Output is discrete (in several aspects)
  - Stream of pixels or images?

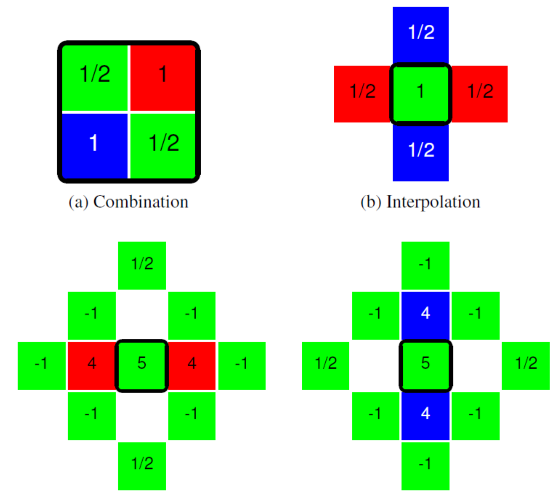
Image sensor

- Color sensors
  - Color Filter Array (CFA)
    - only light of a specific color will excite a photo diod
  - Bayer pattern



Images from Wikipedia

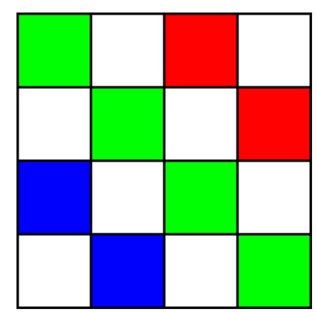
- Color vs monochrome?
  - 5Mp is <u>not</u> 5Mp RGB
  - Demosaicing => computational cost and artifacts
  - Less light-sensitive



(a) Combination (b) Interpolation

(c) Malvar-He-Cutler

Bayer demosaicing,
green pixel - red row



Sparse CFA