Basics of Machine Vision

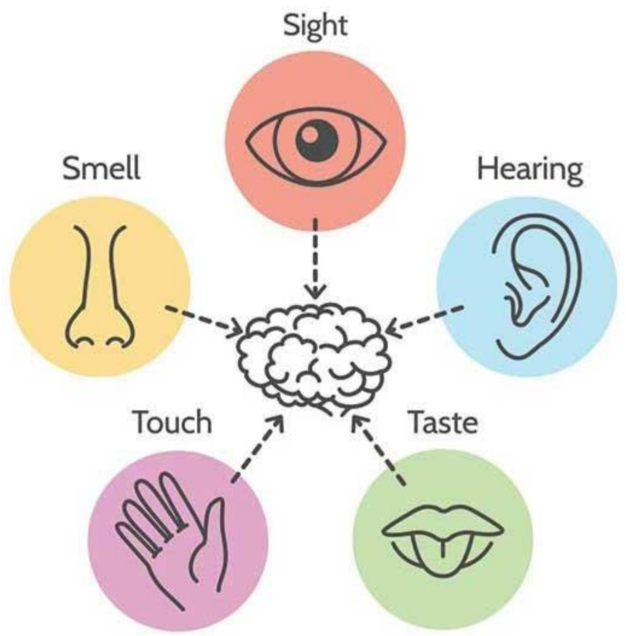
Primož Podržaj

Lecture 01

The motivation for image processing

Human senses:

- Sight or vision.
- Hearing or audition.
- Smell or olfaction.
- Taste or gustation.
- Touch or tactition.



Which one gives the most information? / Which one is the most difficult to live without?

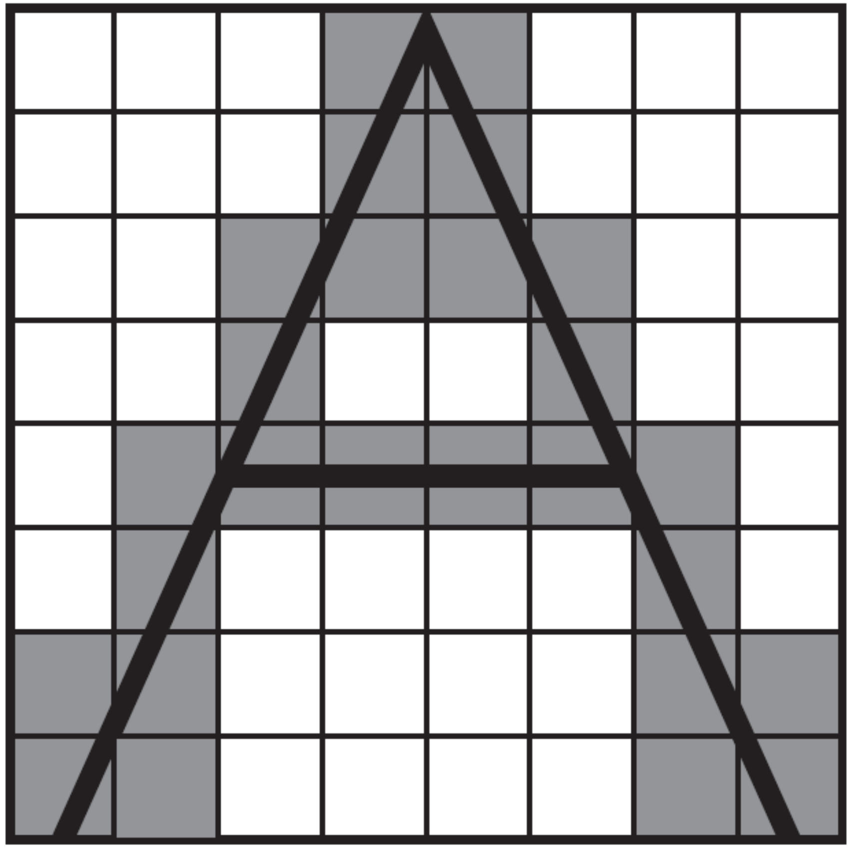
The difference between image processing and computer vision

In image processing, an image is "processed", that is, transformations are applied to an input image and an output image is returned. The transformations can e.g. be "smoothing", "sharpening", "contrasting" and "stretching". The transformation used depends on the context and issue to be solved.

In computer vision, an image or a video is taken as input, and the goal is to understand (including being able to infer something about it) the image and its contents. Computer vision uses image processing algorithms to solve some of its tasks.

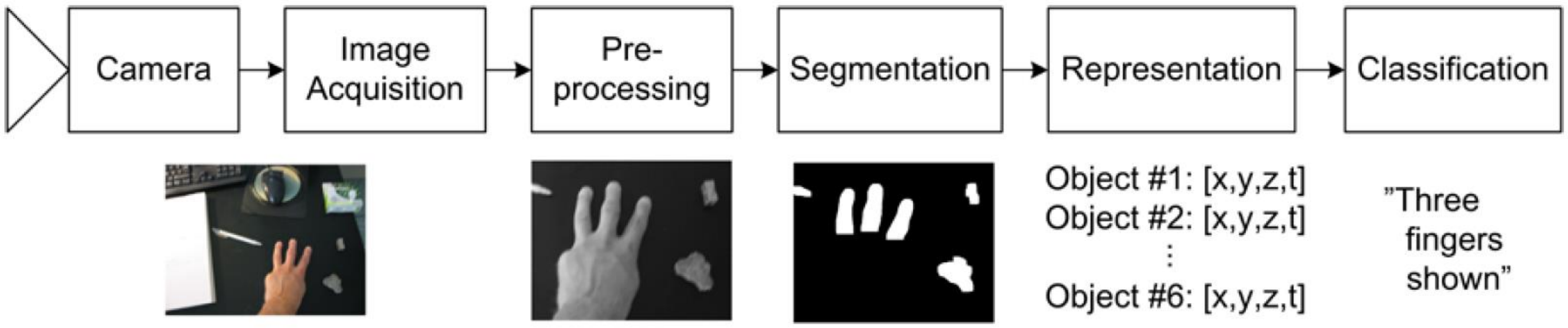
The main difference between these two approaches are the goals (not the methods used). For example, if the goal is to enhance an image for later use, then this may be called image processing. If the goal is to emulate human vision, like object recognition, defect detection or automatic driving, then it may be called computer vision.

The goal of image processing / machine vision

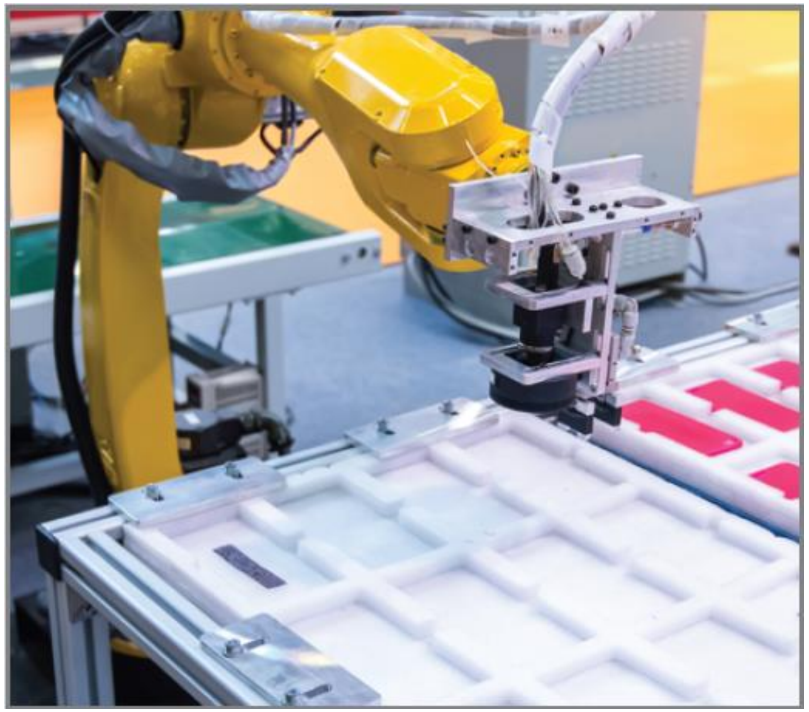


$$x = \begin{bmatrix} 255 & 255 & 255 & 0 & 0 & 255 & 255 & 255 \\ 255 & 255 & 255 & 0 & 0 & 255 & 255 & 255 \\ 255 & 255 & 0 & 0 & 0 & 0 & 255 & 255 \\ 255 & 255 & 0 & 255 & 255 & 0 & 255 & 255 \\ 255 & 0 & 0 & 0 & 0 & 0 & 0 & 255 \\ 255 & 0 & 255 & 255 & 255 & 255 & 0 & 255 \\ 0 & 0 & 255 & 255 & 255 & 255 & 0 & 0 \\ 0 & 0 & 255 & 255 & 255 & 255 & 0 & 0 \end{bmatrix}$$

Typical steps in computer vision

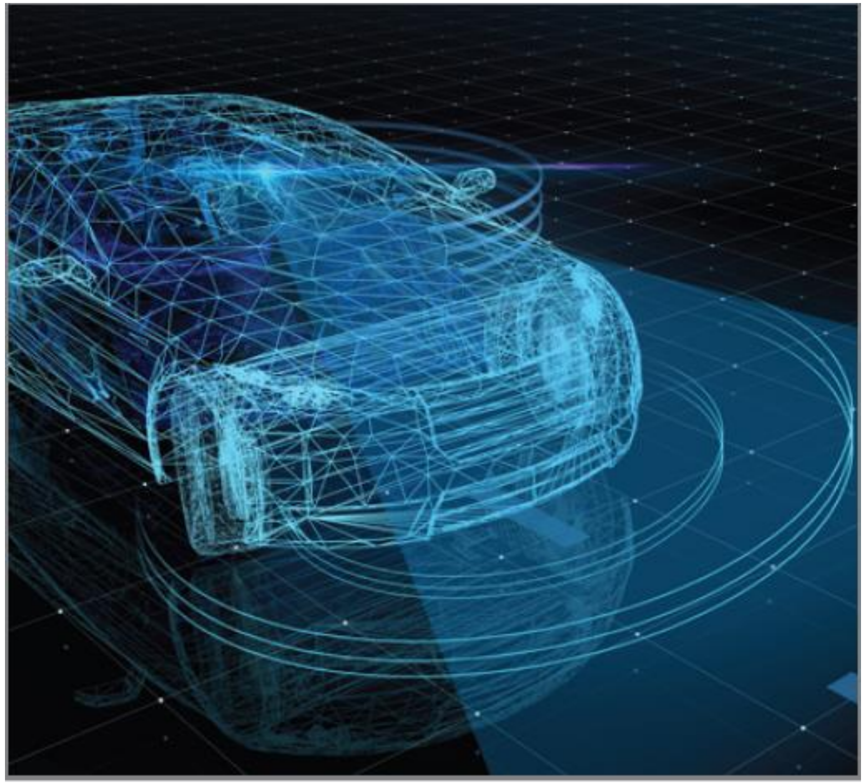


Applications 1 – Factory automation



The use of controllers, algorithms, and sensors to automate repetitive tasks and reduce human oversight. Commonly automated tasks include sorting, inspection, and defect detection. In general, when thinking of “machine vision”, factory automation is what springs to mind.

Applications 2 – Autonomous Systems

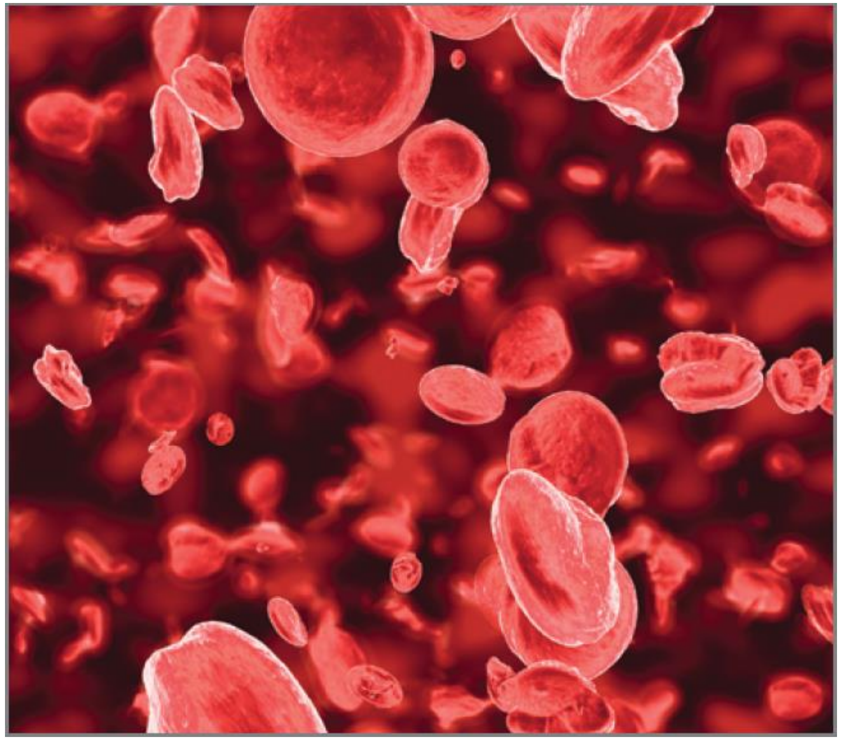


Autonomous means having the ability to self-govern. Common autonomous systems include self-driving cars and trucks, flying taxis, agriculture or farming robots, and delivery robotics. Vision systems are an incredibly important piece for the future of autonomous systems.

Applications 3 – Logistics

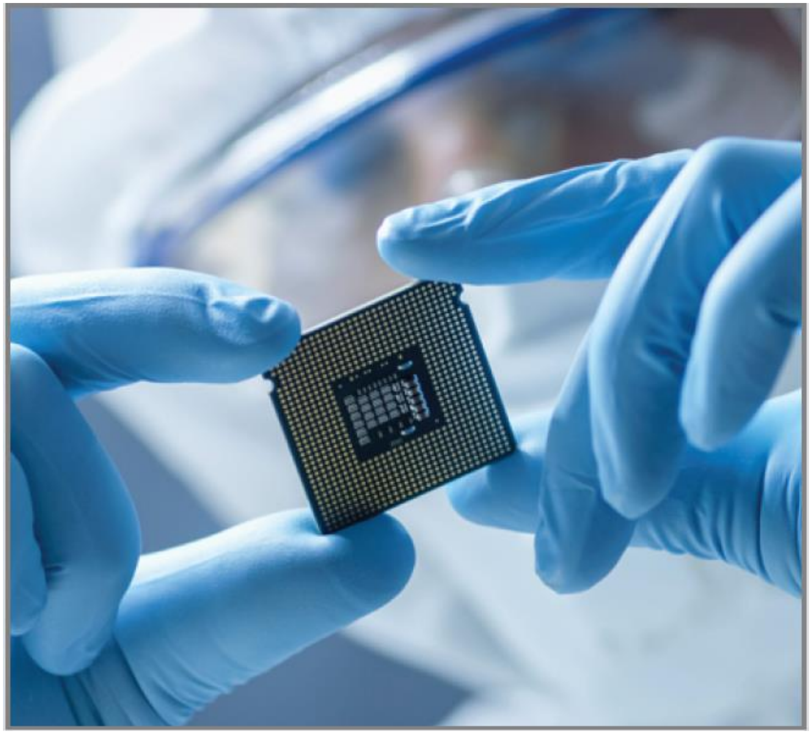
Logistic processes often use robots for automated warehousing. Robots perform OCR or scan barcodes on products to rapidly identify products on shelves or packaged and ready to be shipped.

Applications 4 – Life Sciences



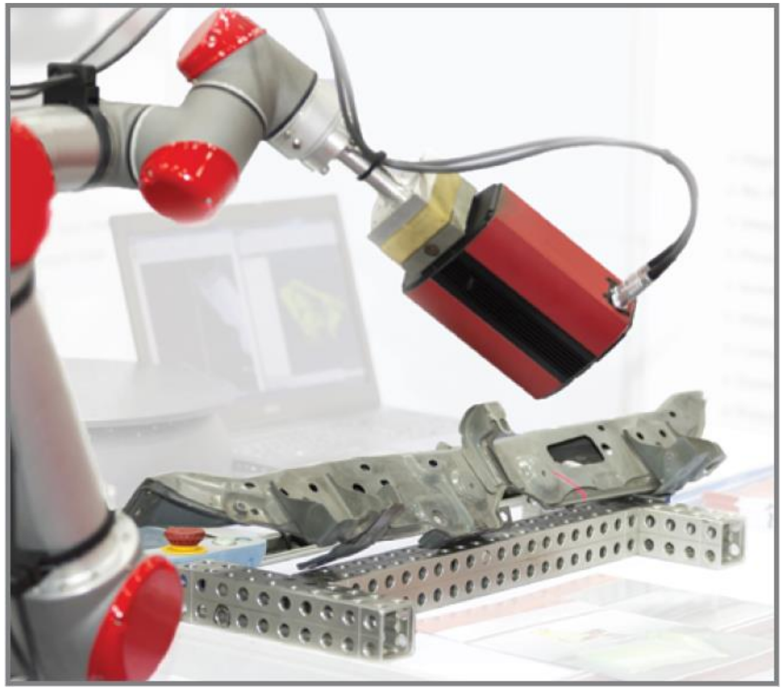
Life sciences include fields related to biology, medicine, physiology, and much more. Besides Xray imaging and MRIs, this space also uses a wide range of imaging techniques like microscopy and special labelling to view, count, sort, and perform other cytometry methods on cells. Photoplethysmography (rPPT)

Applications 5 – Electronics and Semiconductors



More circuitry can be integrated on semiconductors than ever before and flat panel displays have extremely high resolutions. To manufacture such complex devices, electronics and displays must be inspected for chip placement and defects at very high resolution.

Applications 6 – Metrology

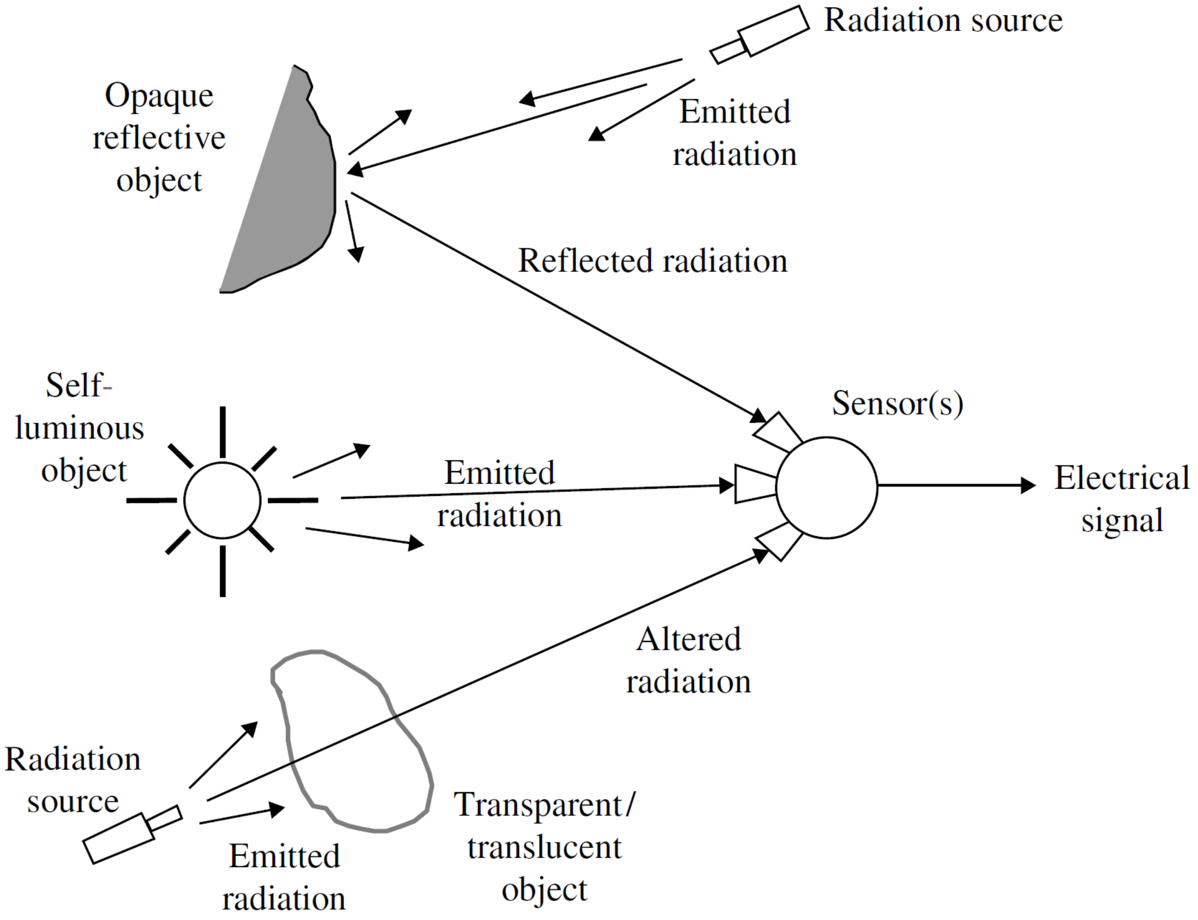


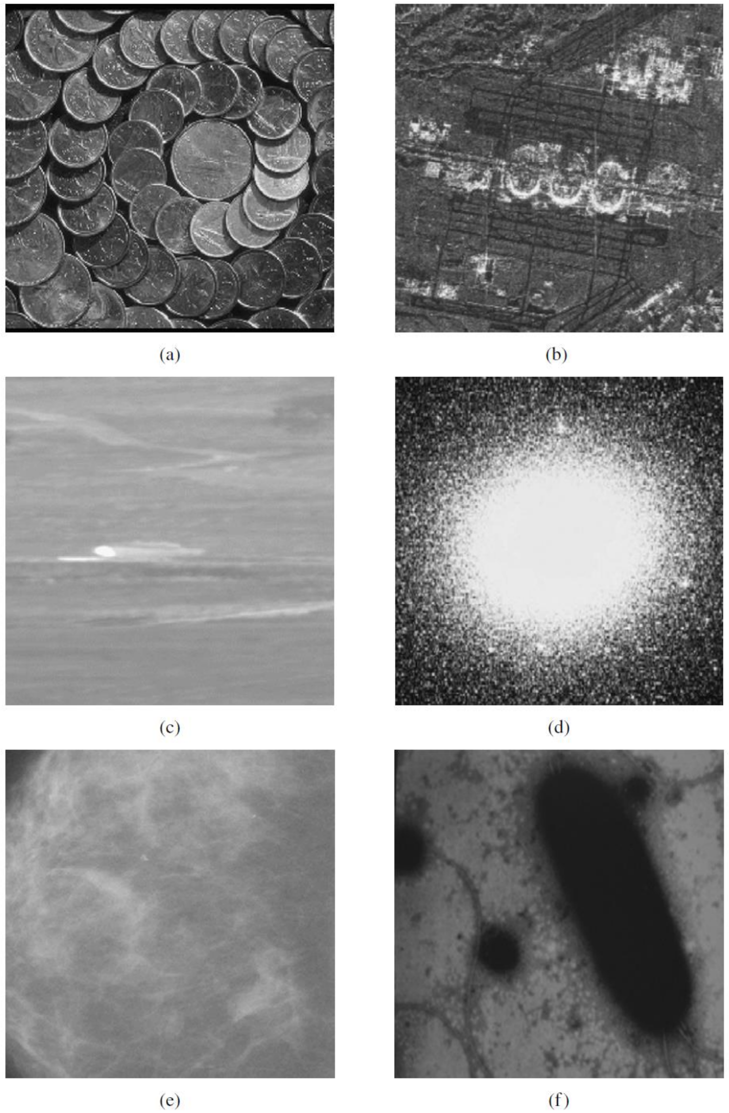
Information about an inspection sample like a characterstic dimension or colors must be measured with repeatable accuracy and reliability. Applications that require accurate measurement include time-of-flight imaging, scheimpflug scanning, 3D imaging, and LIDAR imaging.

Applications 7 – Agriculture

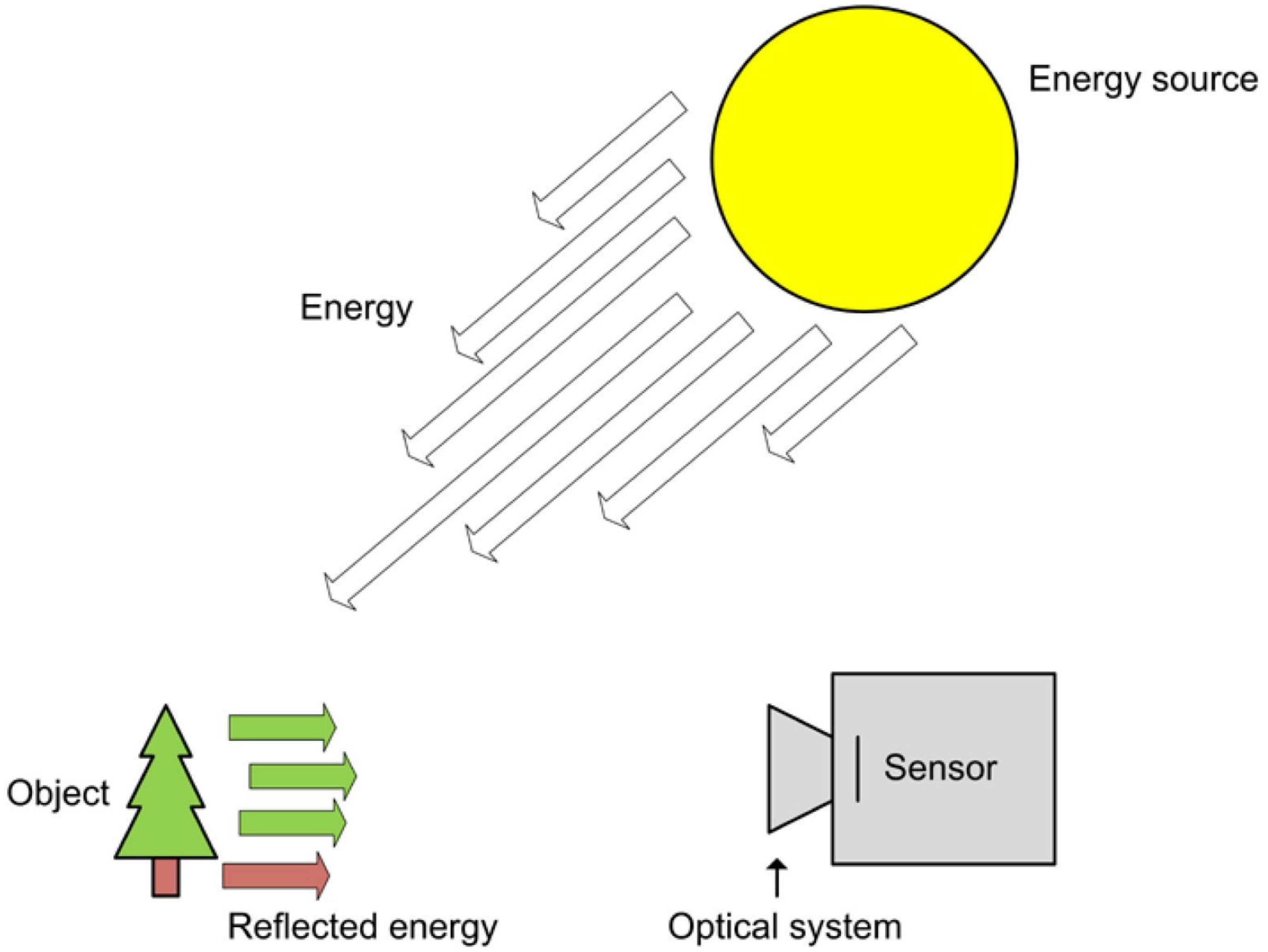
Various types of agriculture related task can be successfully implemented with the help of visual inspection (Colorado beetle or hornet nest detection, soil humidity measurement, automated plant picking, spruce cone picking, automated root system characterization, varroa fighting).

Recording the various types of interaction of radiation with matter

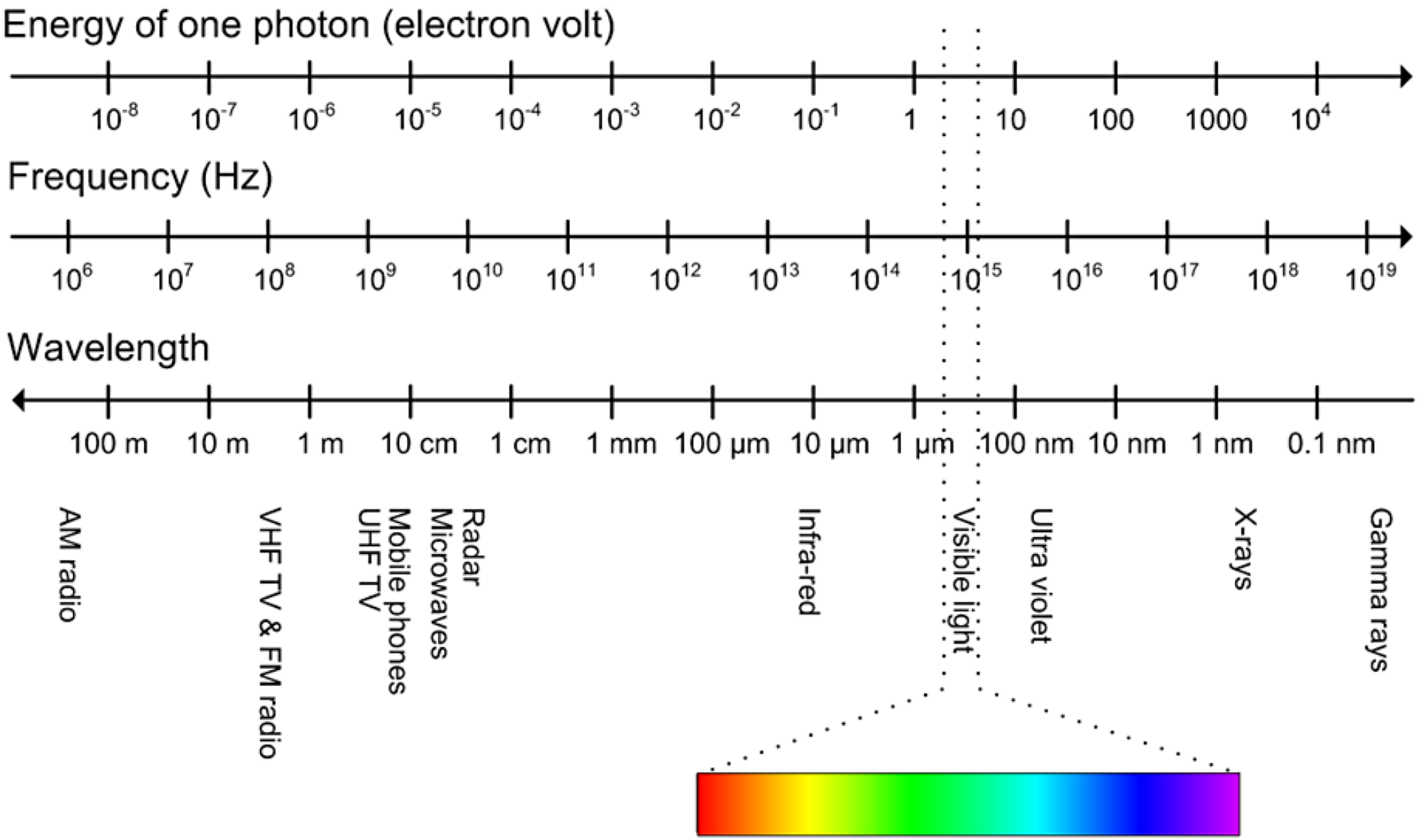




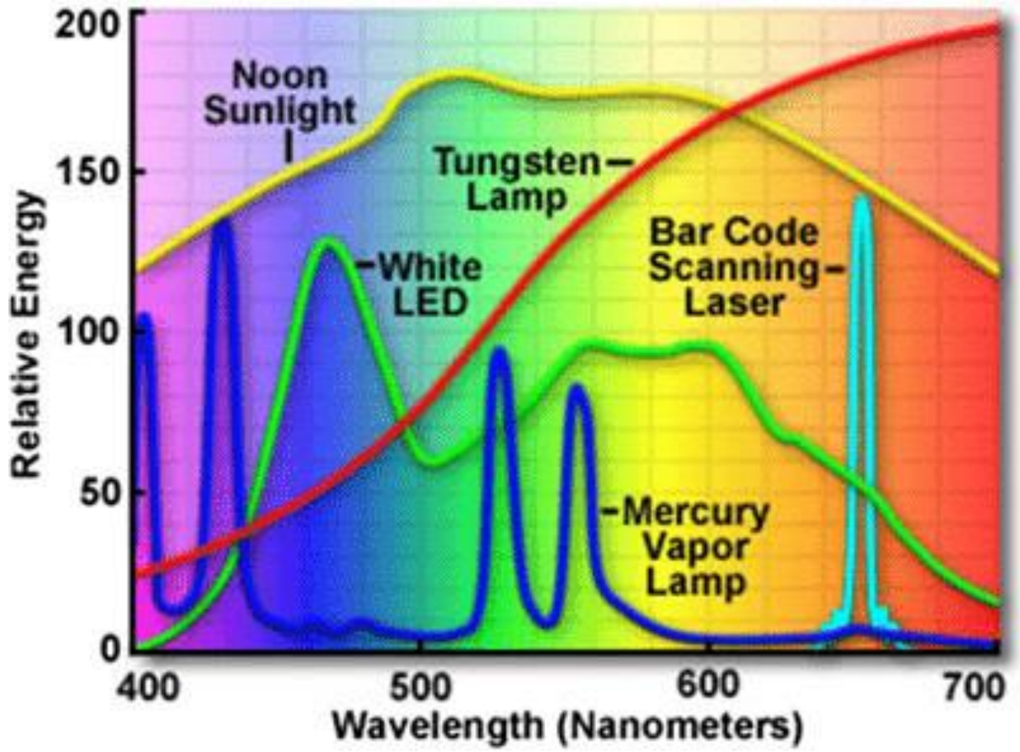
Basics of image acquisition



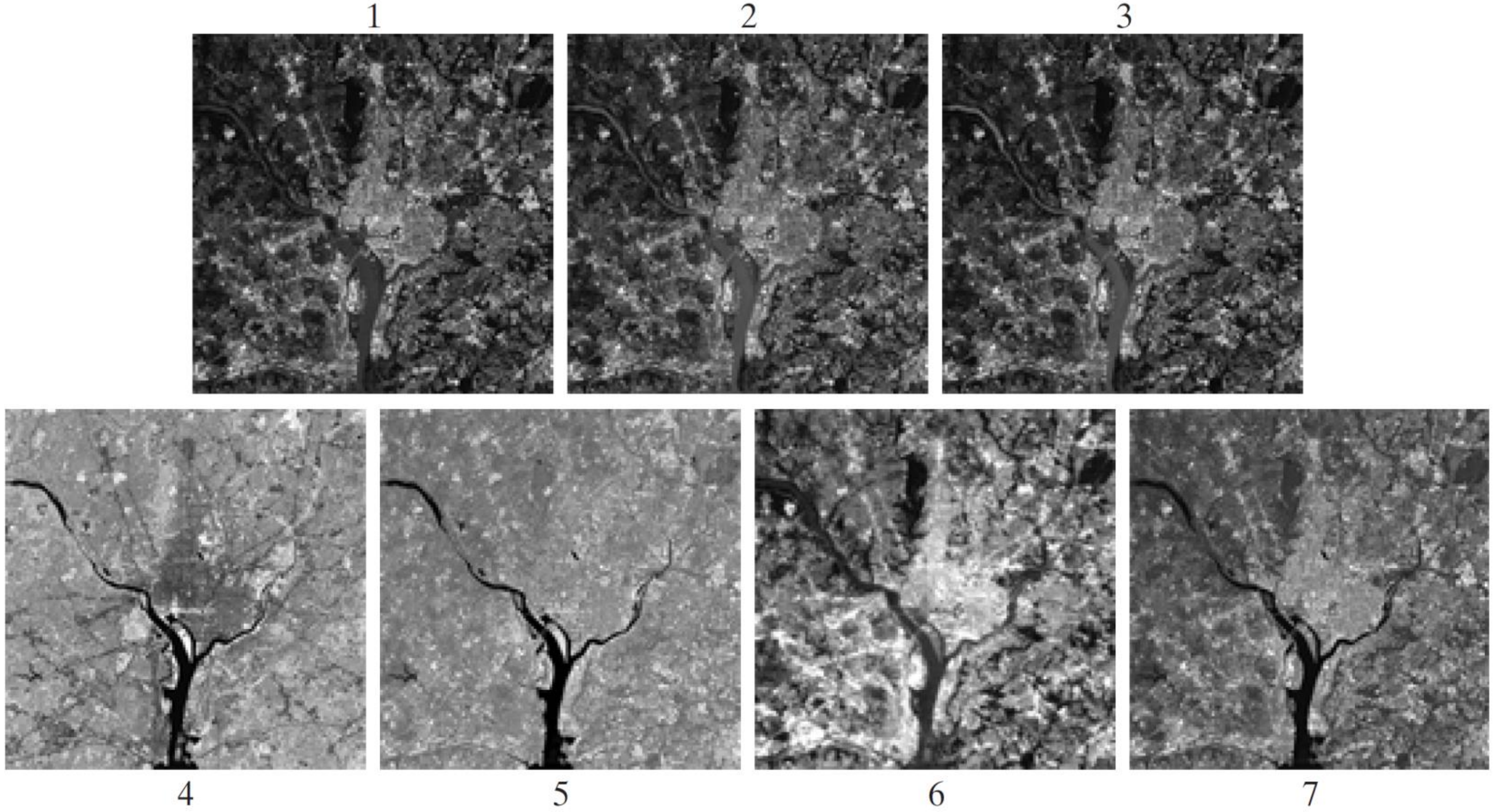
Different parts of EM spectrum used in image processing



Frequency spectrum of different light sources



Images at different frequencies



Charecteristic application

Band No.	Name	Wavelength (μm)	Characteristics and Uses
1	Visible blue	0.45–0.52	Maximum water penetration
2	Visible green	0.52–0.60	Good for measuring plant vigor
3	Visible red	0.63–0.69	Vegetation discrimination
4	Near infrared	0.76–0.90	Biomass and shoreline mapping
5	Middle infrared	1.55–1.75	Moisture content of soil and vegetation
6	Thermal infrared	10.4–12.5	Soil moisture; thermal mapping
7	Middle infrared	2.08–2.35	Mineral mapping

Types of cameras

- Monochromatic
- Color (RGB)
- Multispectral (3 to 15 spectral bands)
- Hyperspectral (~100 spectral bands or more)

Imaging sensor (line vs array)

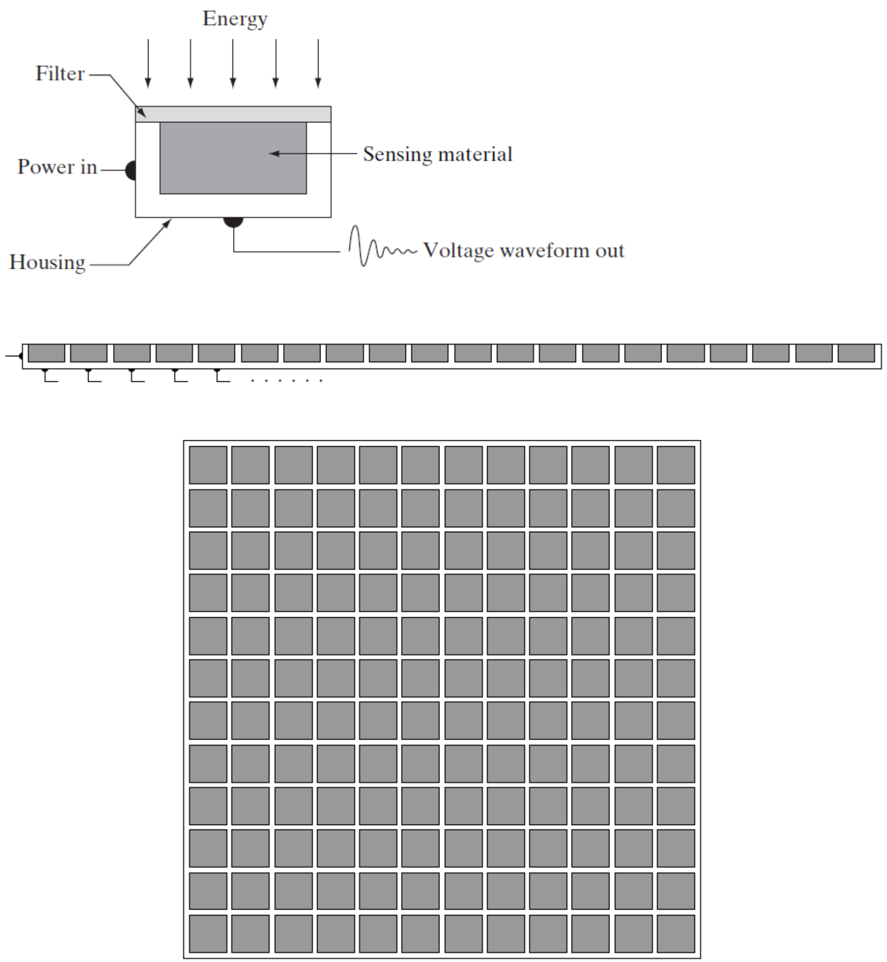


Image Acquisition Using a Single Sensor

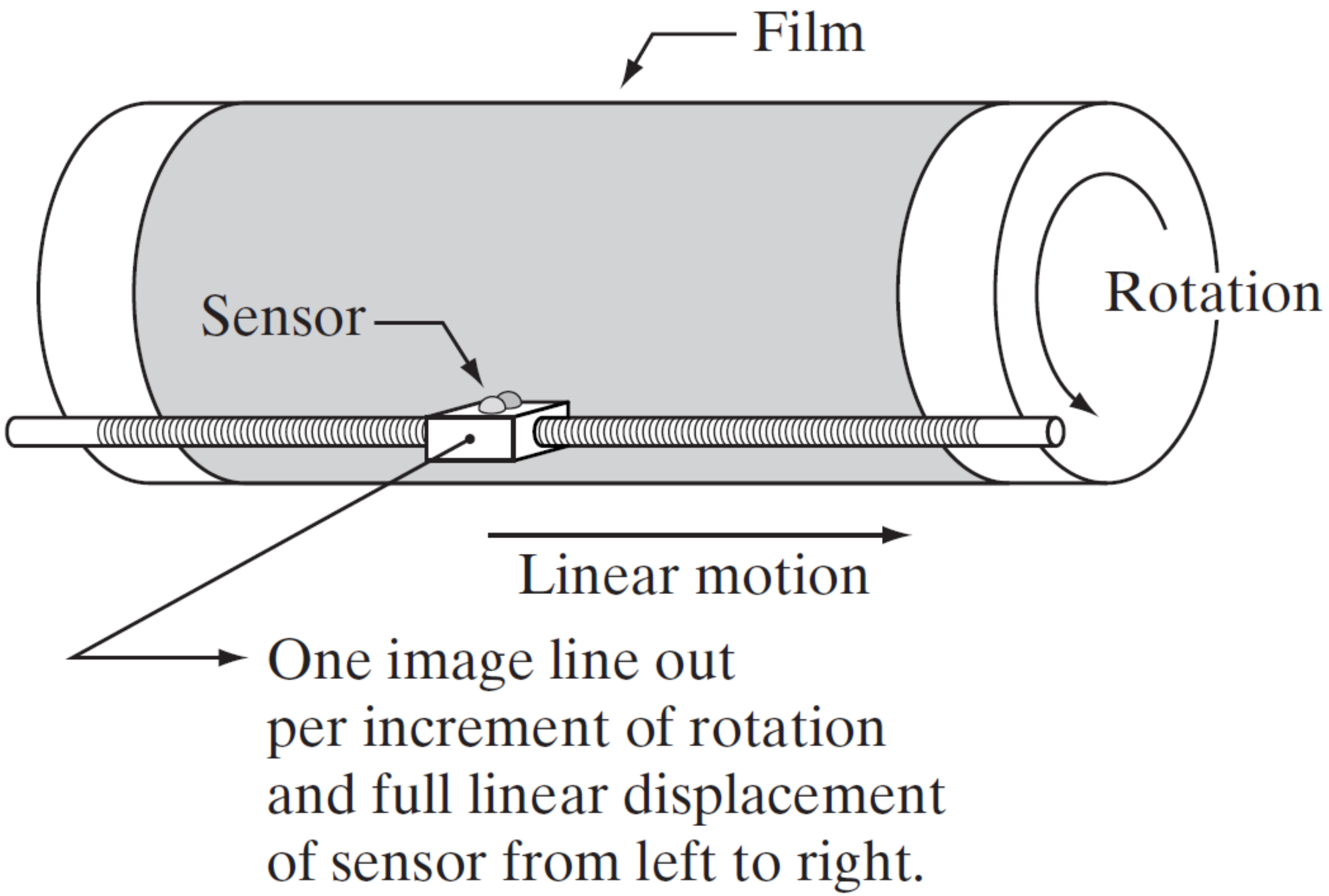
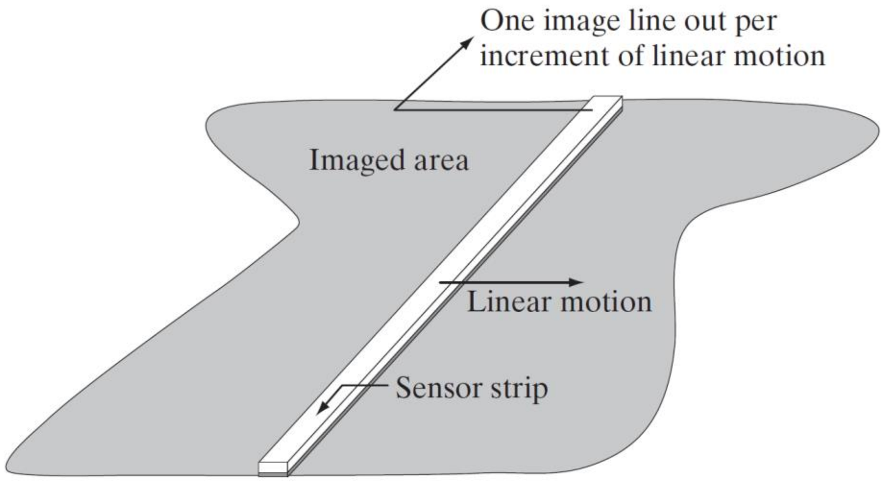


Image Acquisition Using Sensor Strips



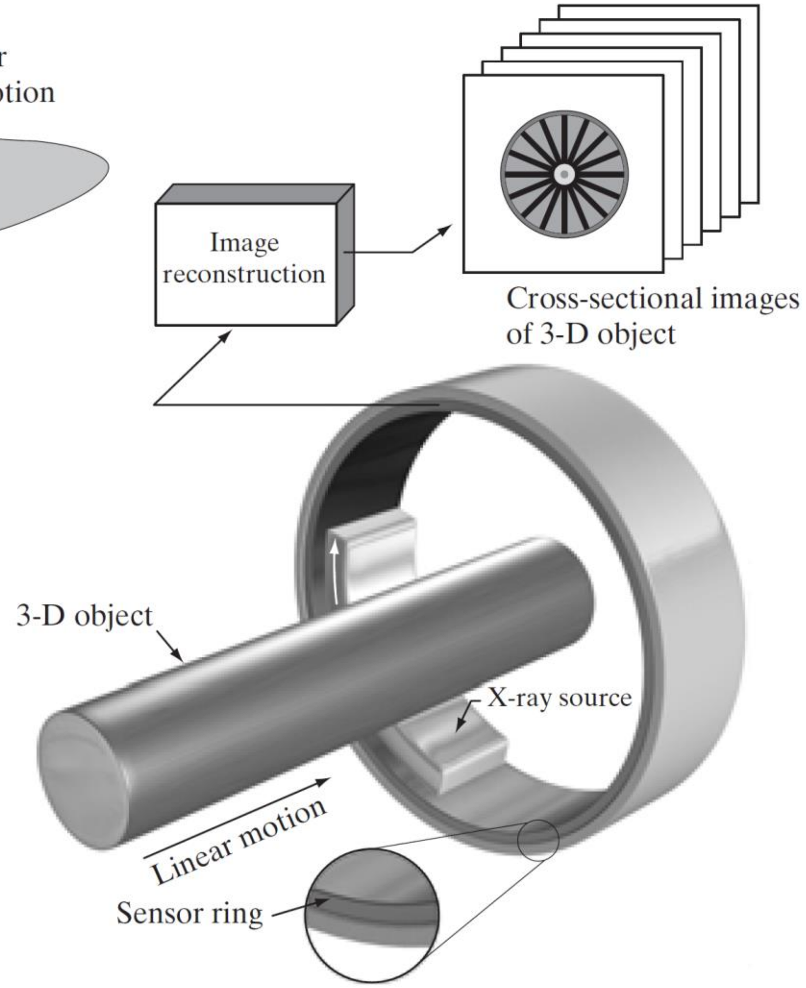


Image Acquisition Using Sensor Arrays

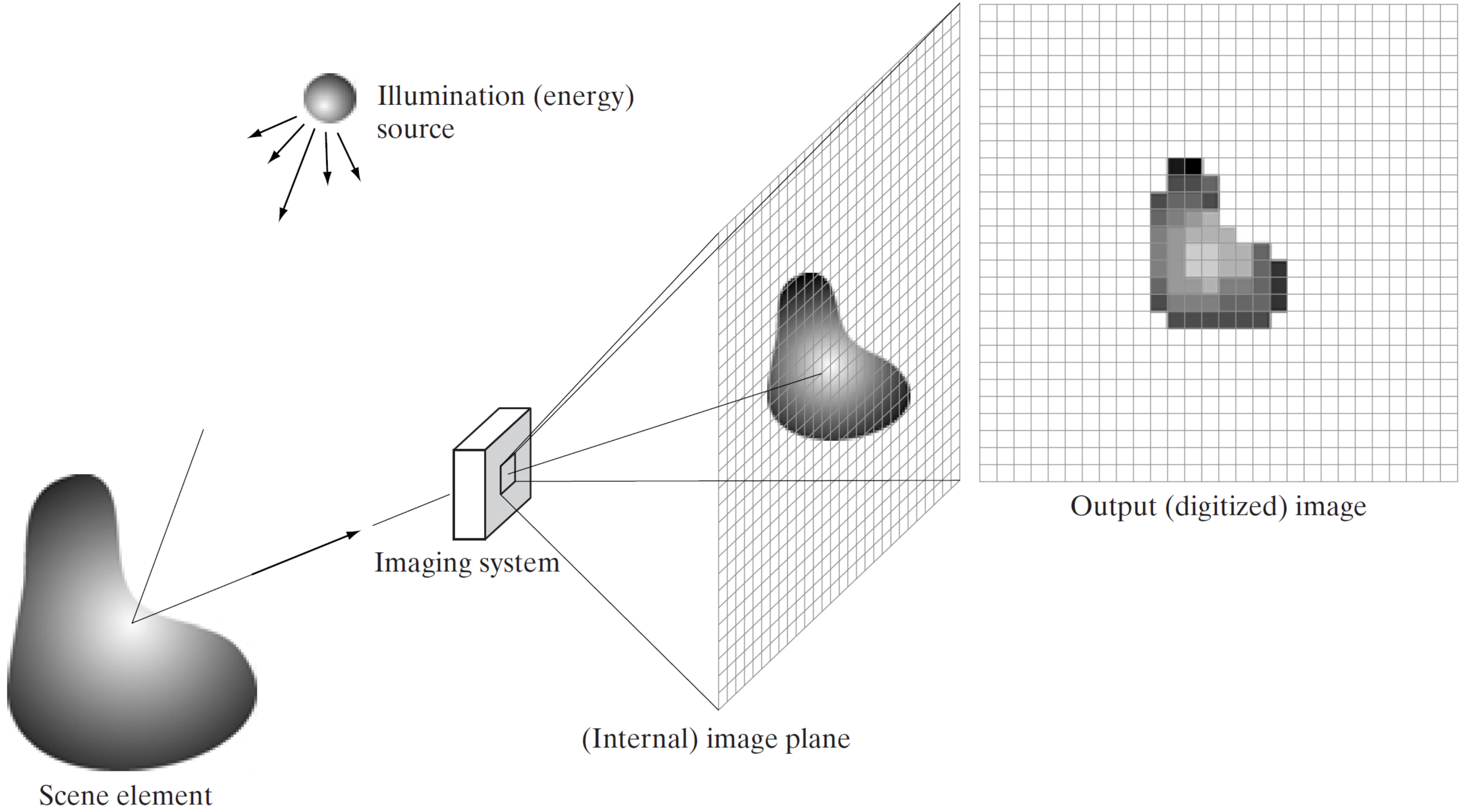
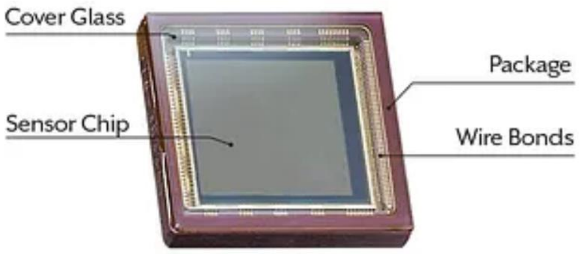


Image sensor



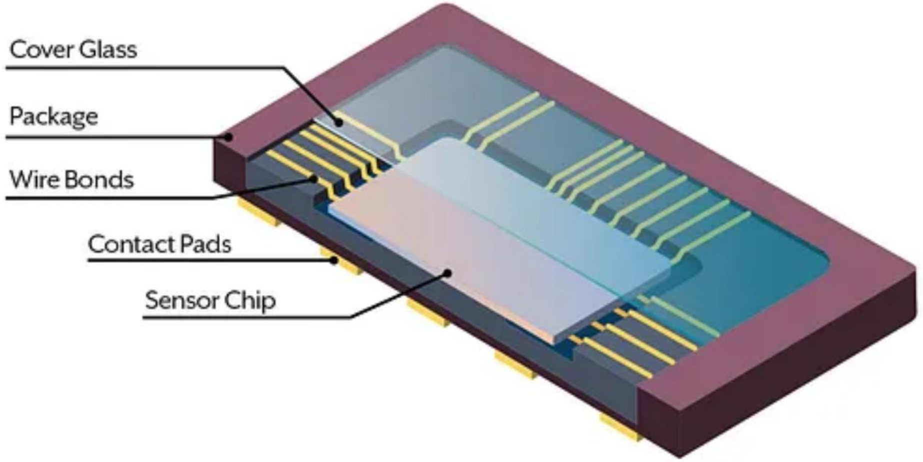
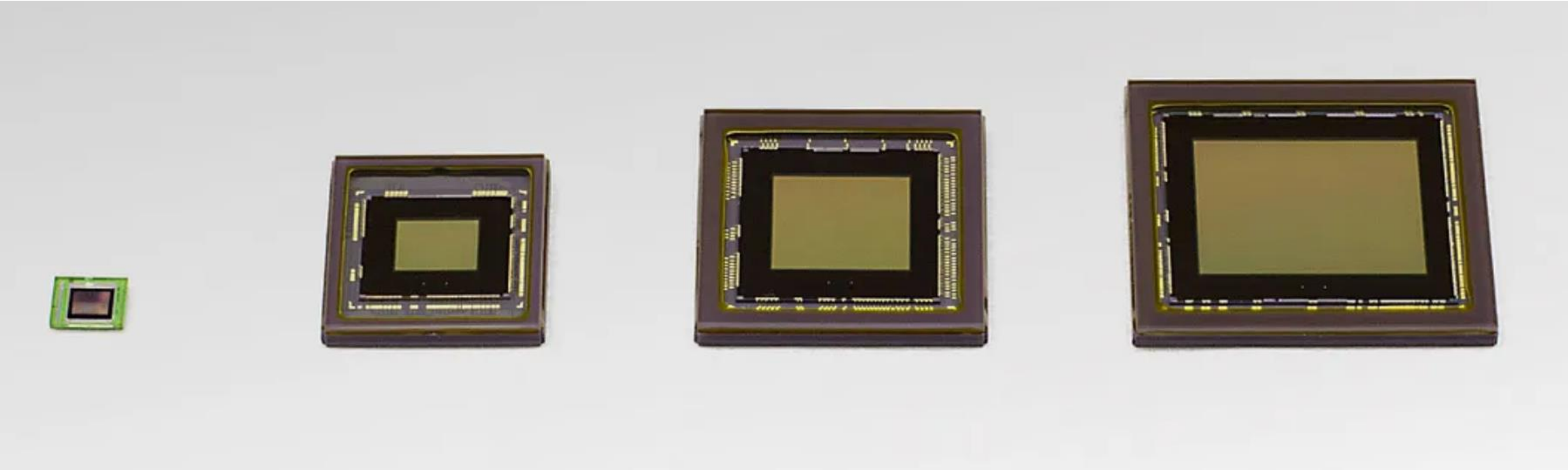


Image Sensor Format (Size)



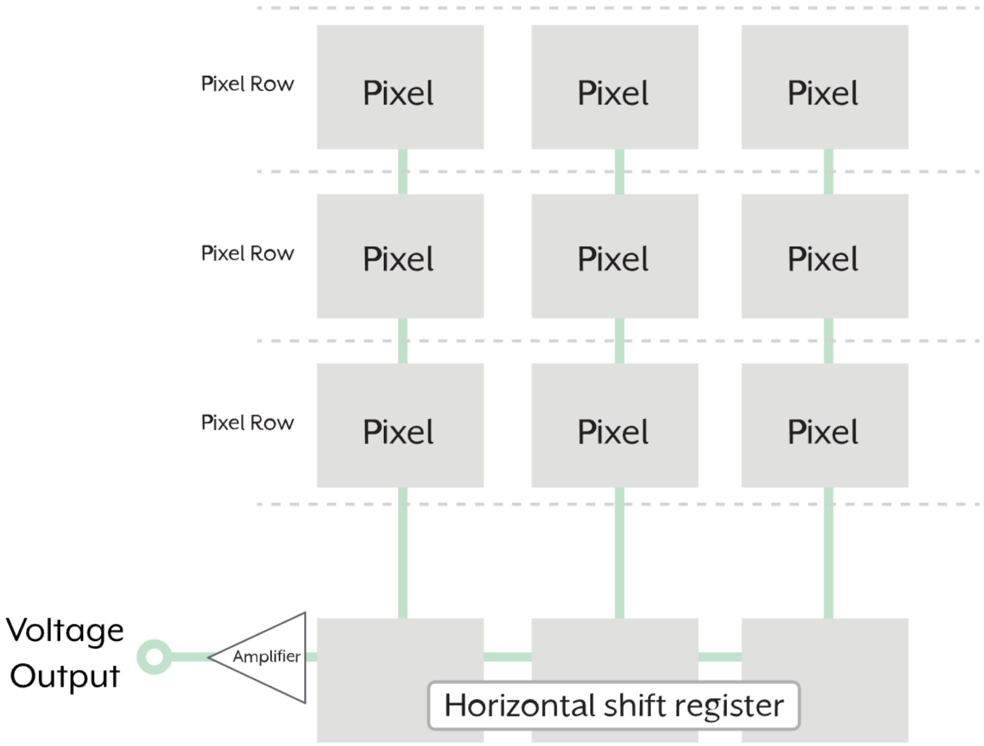
Sensor Format Sizes (Left to Right): 1/6″, 1/3″, 2/3″, 1″

Types of sensors

- CCD (Charged Couple Device)
- CMOS (Complementary Metal-Oxide Semiconductor)

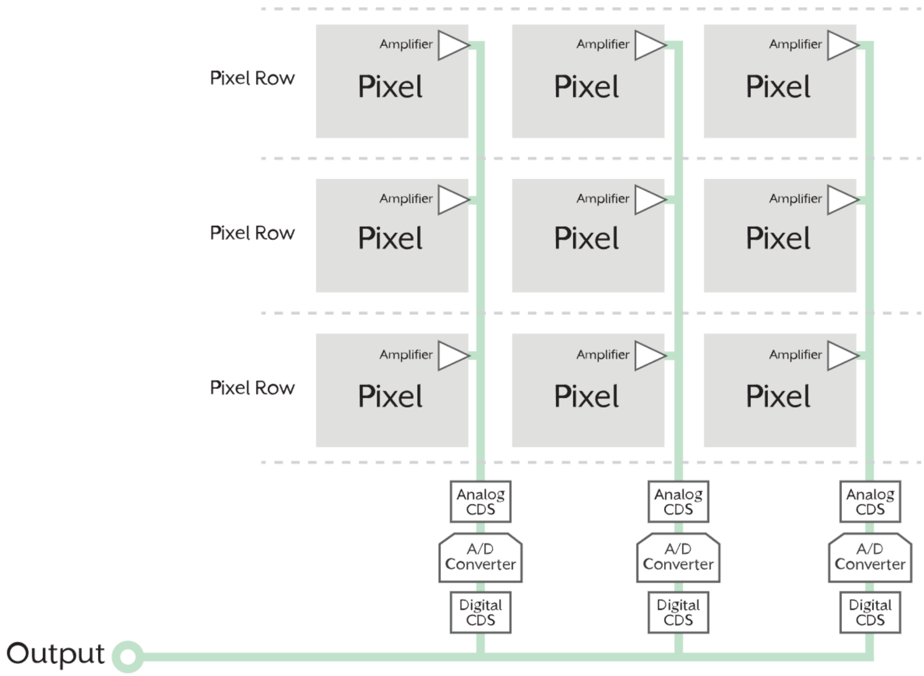
CCD

- CCD sensors (Charged Couple Device) start and stop exposure for all pixels at the same time. This is known as global shutter.

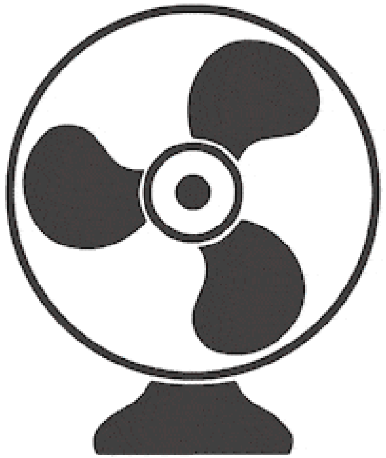


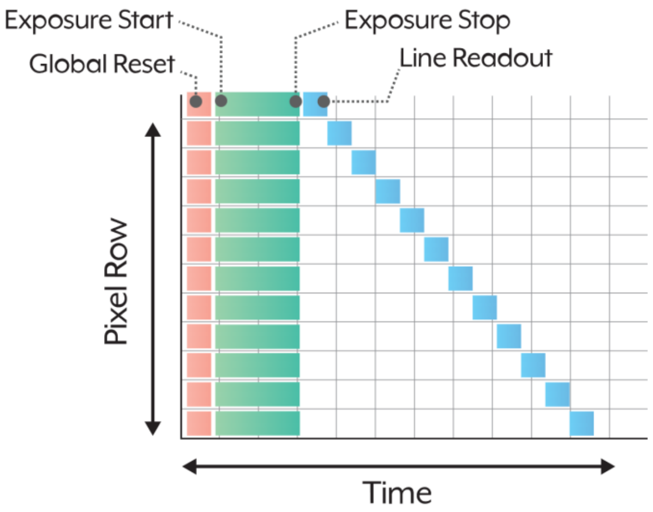
CMOS

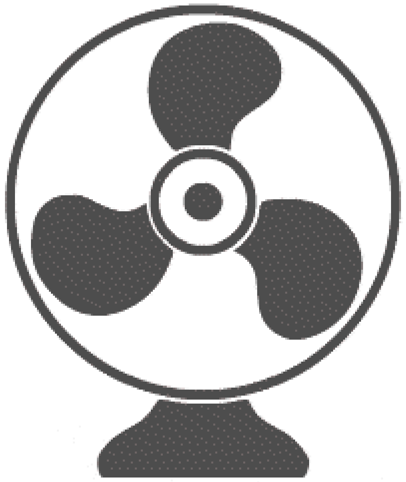
- In the past, CMOS sensors (Complementary Metal-Oxide Semiconductor) were only able to start and stop exposure one pixel row at a time, which is known as rolling shutter. This has changed over time, with many global shutter CMOS sensors now available in the market.

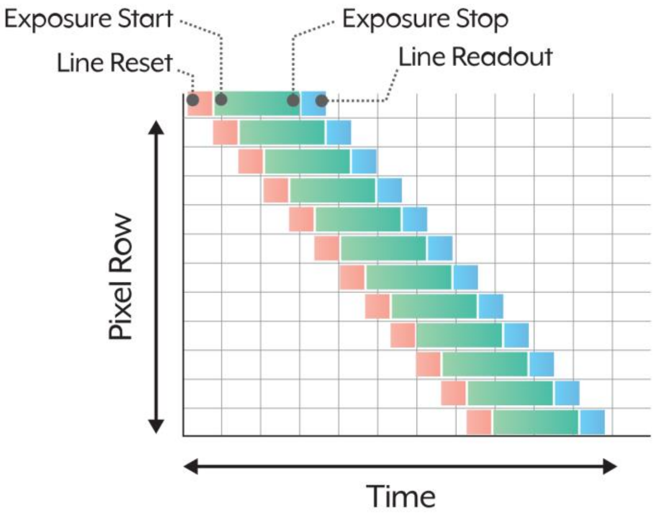


Global vs. rolling shutter

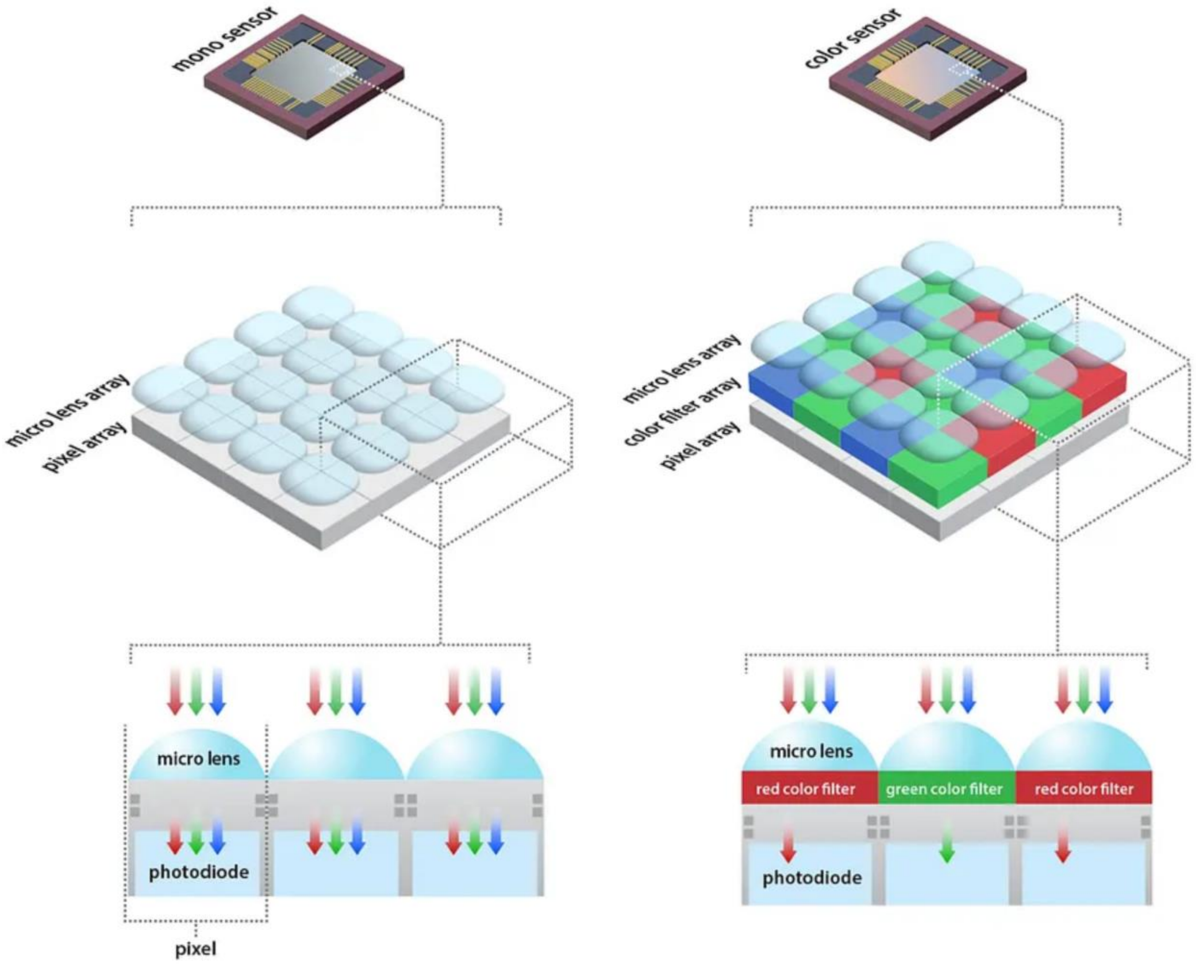




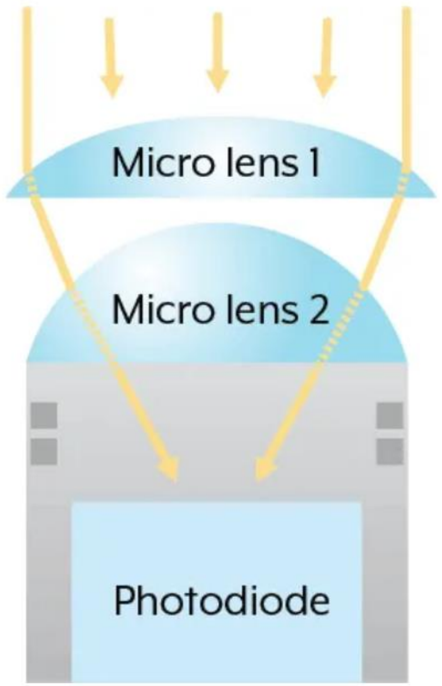




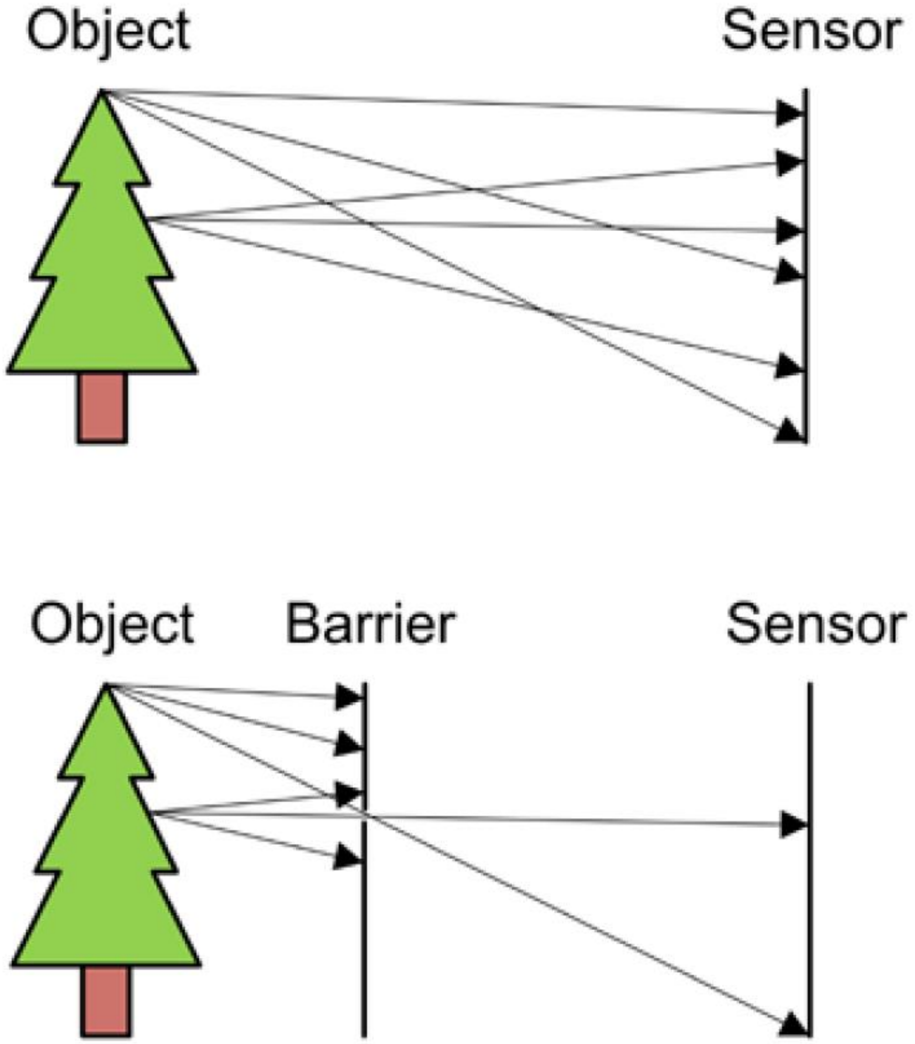
Mono and color sensors

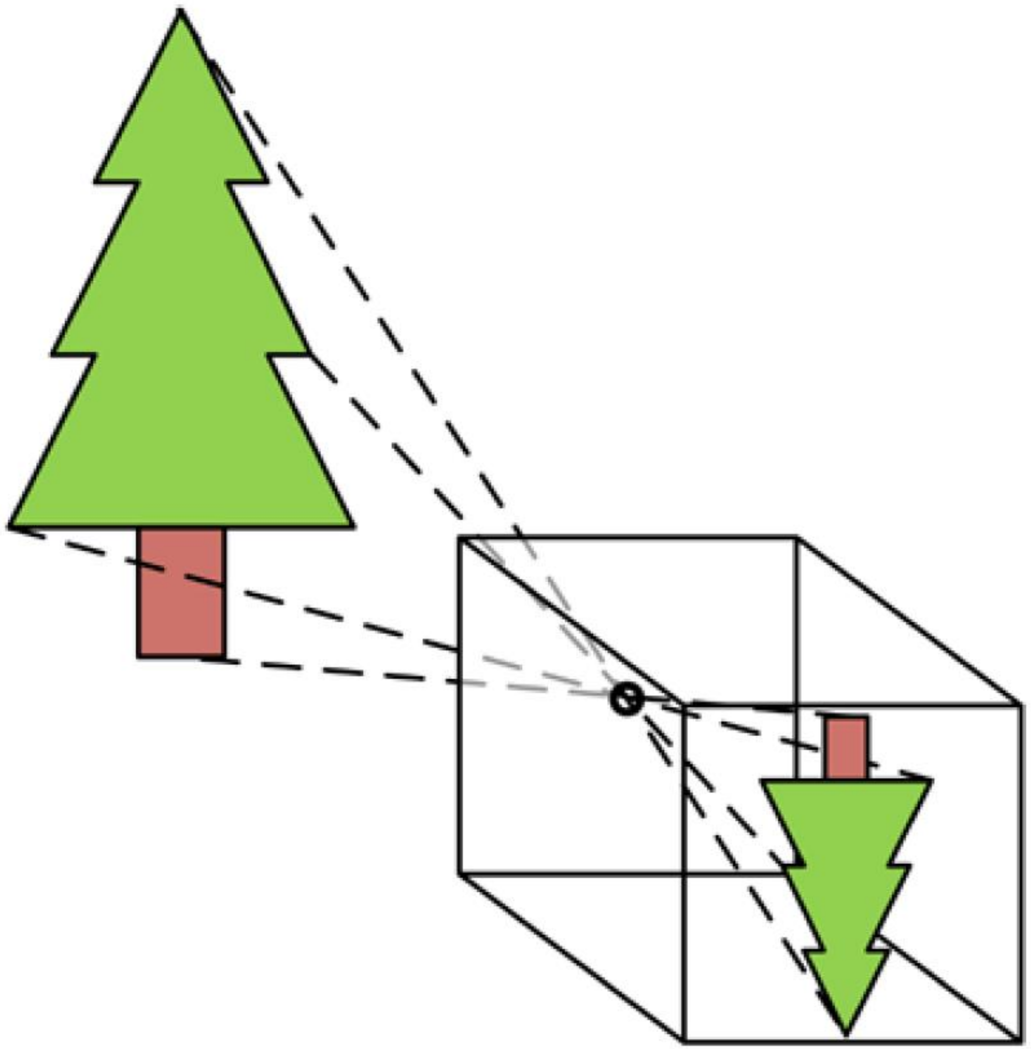


Solution for sensors with very small pixel sizes



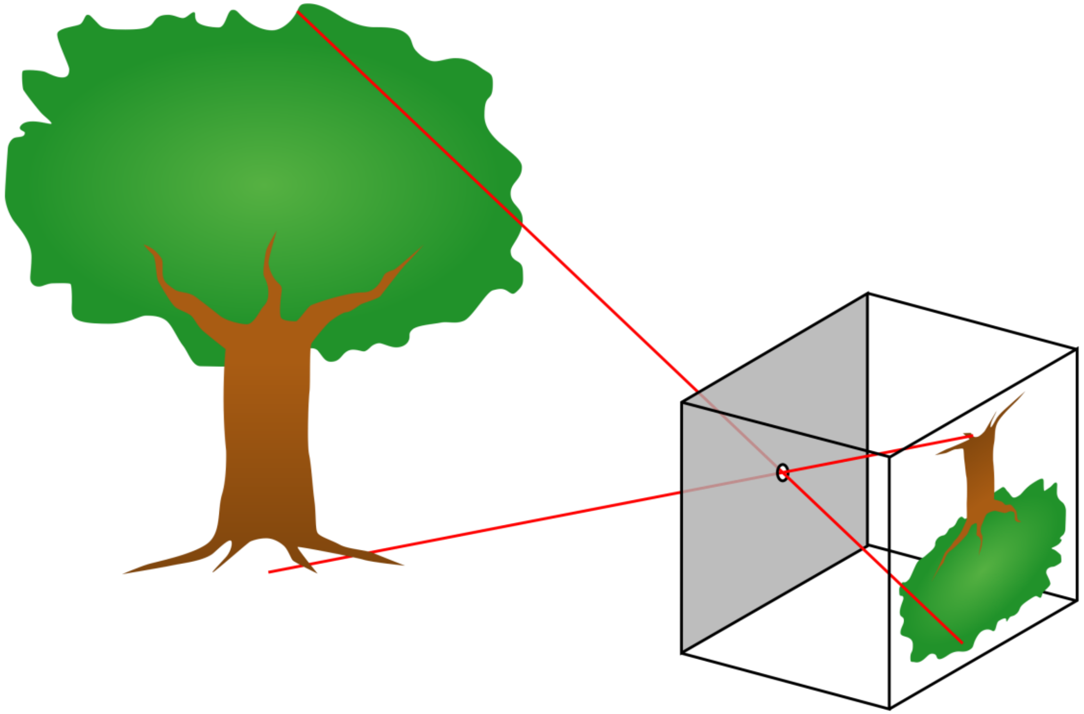
Basics of the optical system



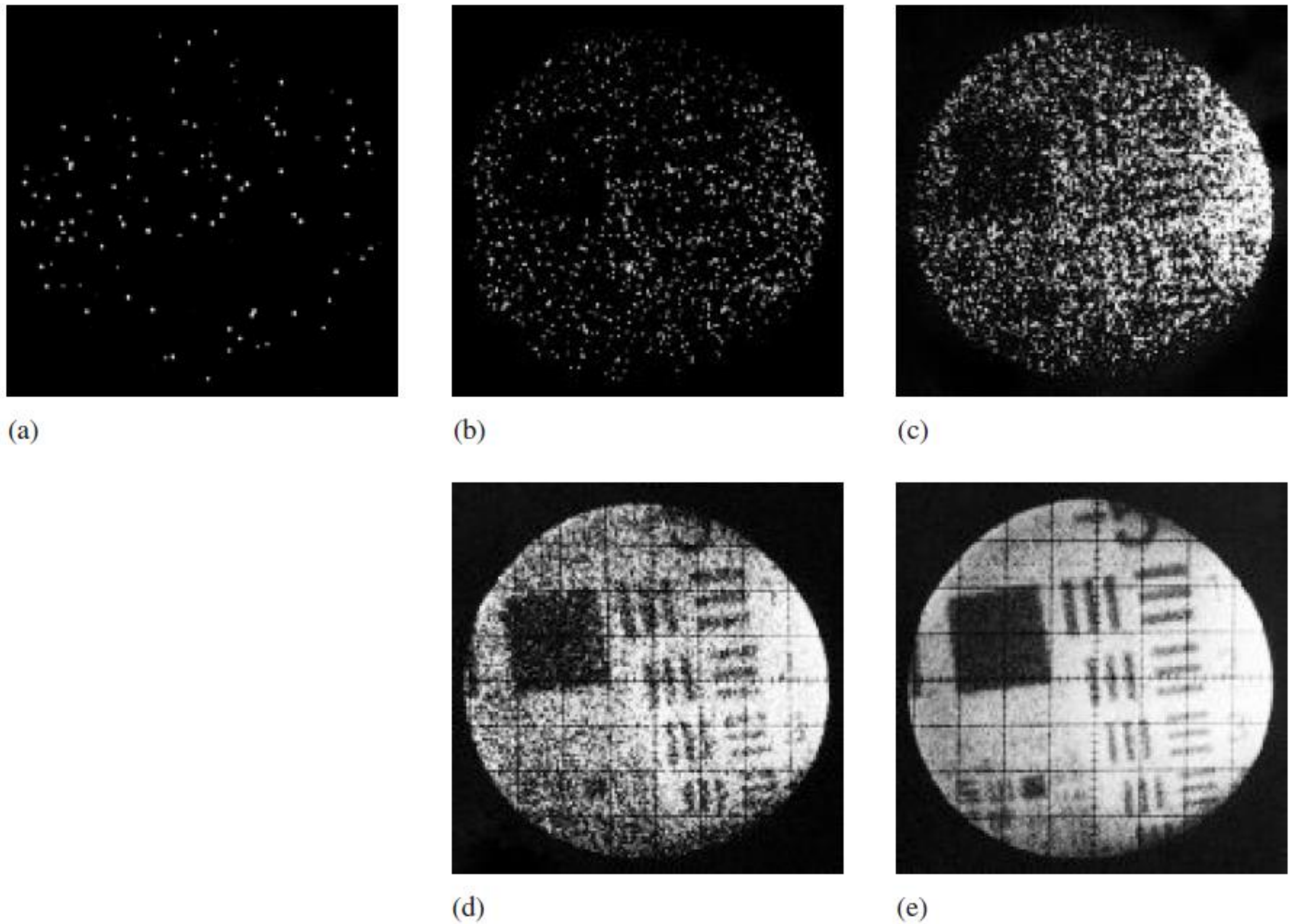


Pinhole camera

Price 330EUR



Why optics is needed?



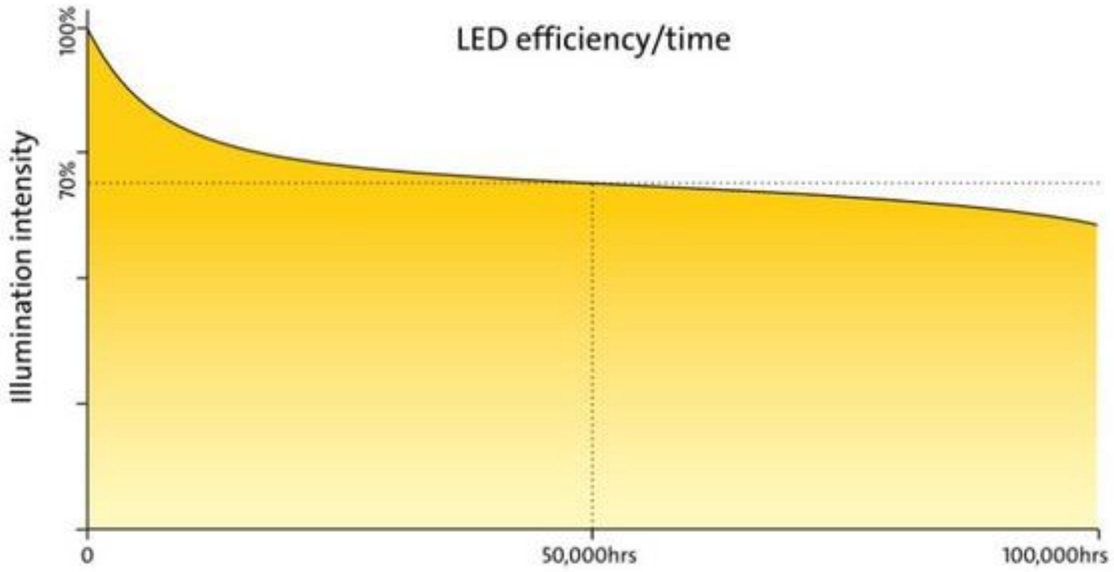
The exposure times were
(a) 8 ms, (b) 125 ms, (c) 1 s, (d) 10 s,
and (e) 100 s.

The value read at a certain pixel

- $f(x, y)=i(x, y)r(x, y)$

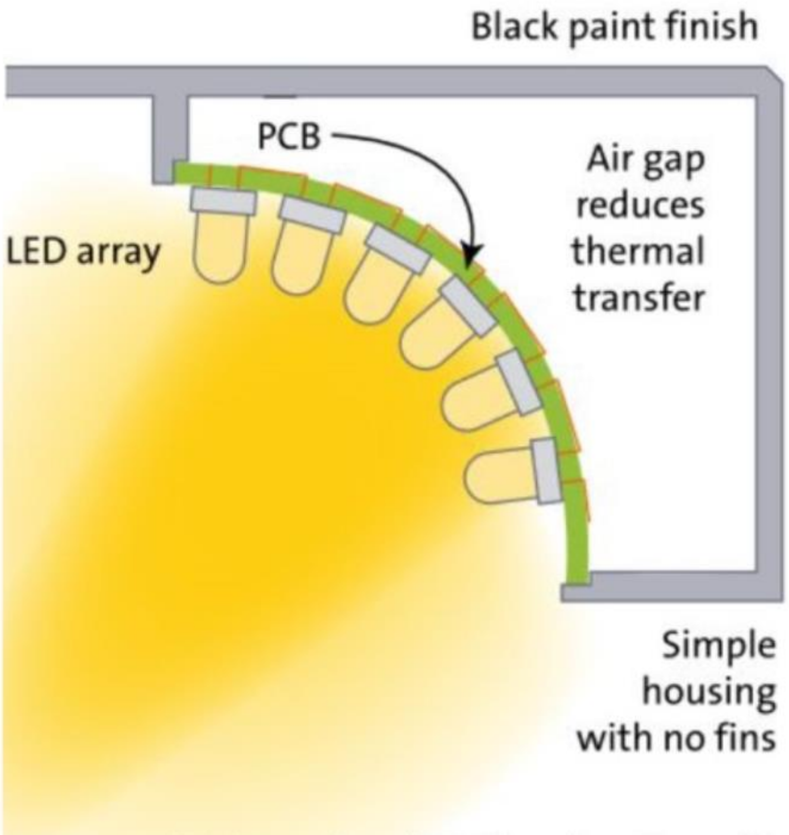
- $f(x, y)$ ………… proportional to energy collected at a certain pixel
- $i(x, y)$ ………… illumination
- $r(x, y)$ ………… reflectance

LED efficiency



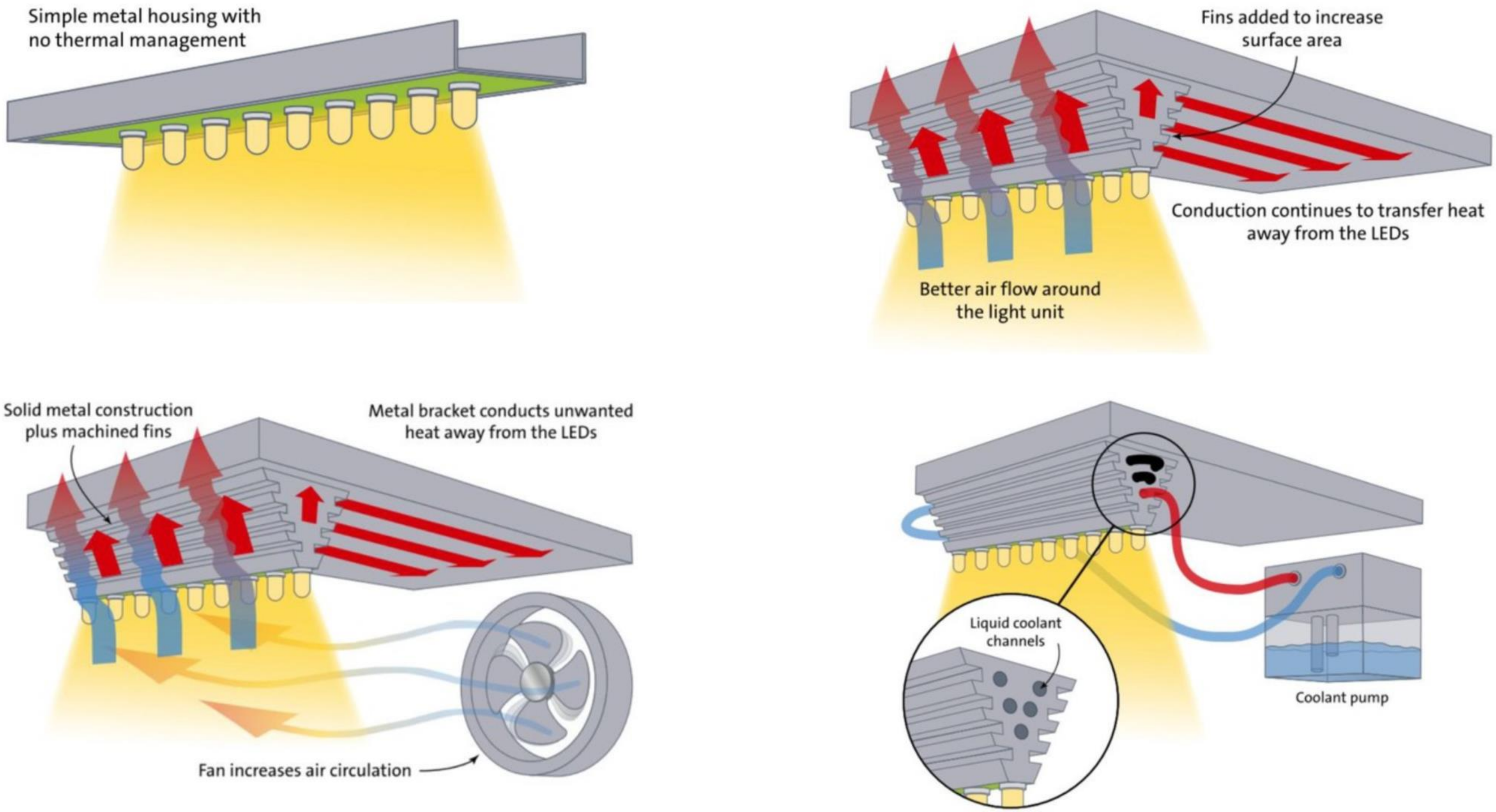
An LED's life span is considered to be the point at which the output intensity drops below 70%.

Heat dissipation problem

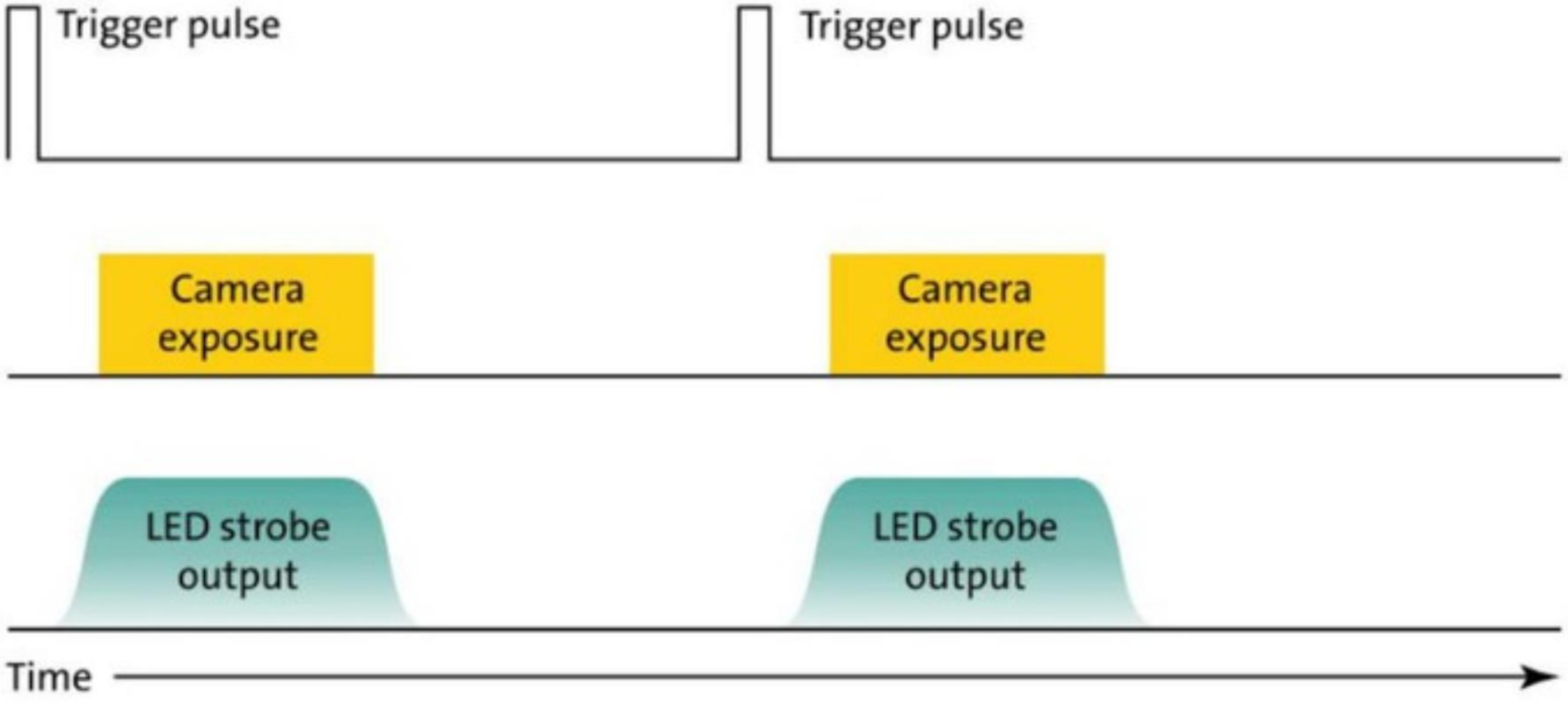


Design of an LED illumination with bad heat dissipation over the housing

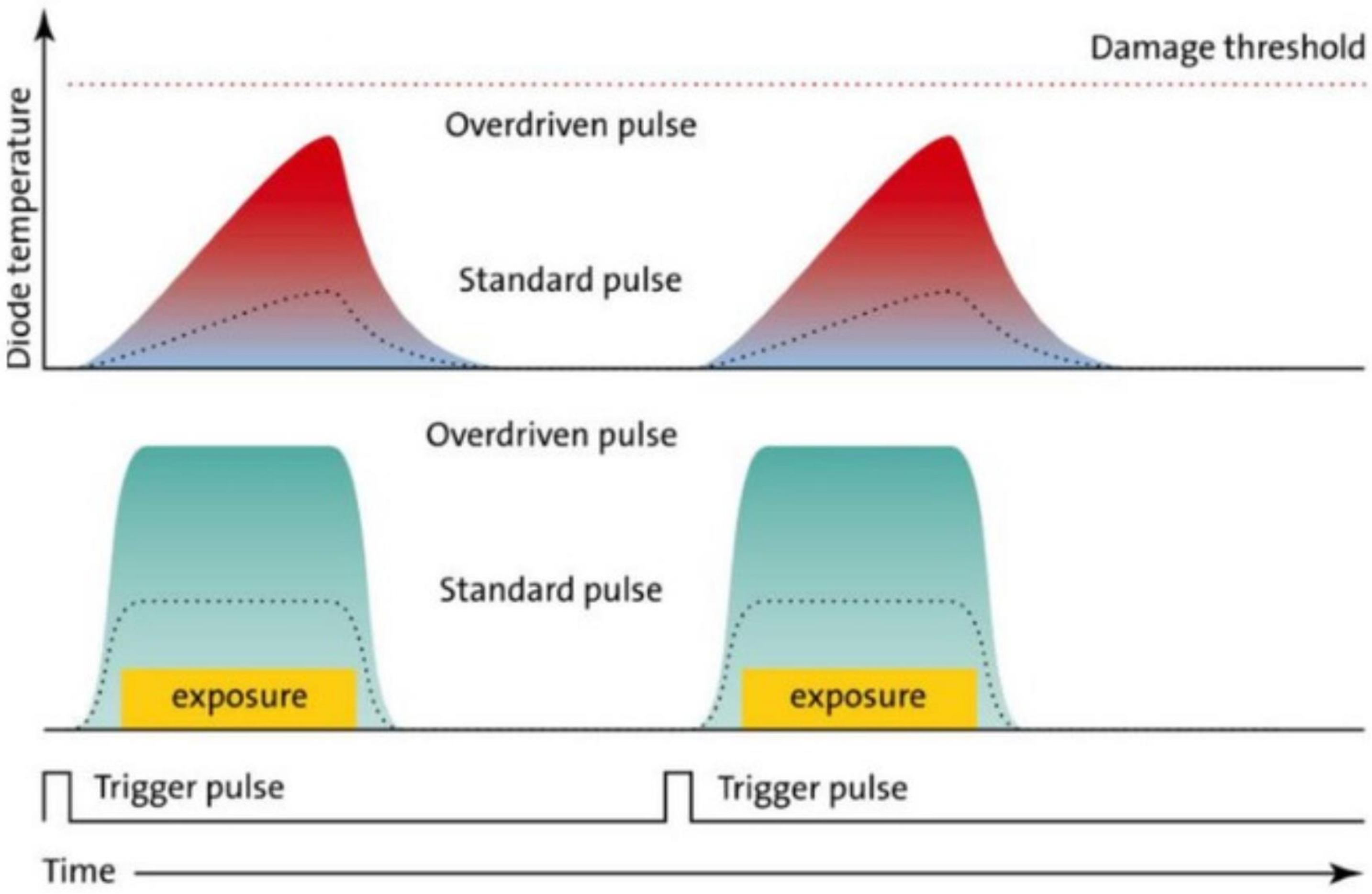
Solutions for better heat dissipation



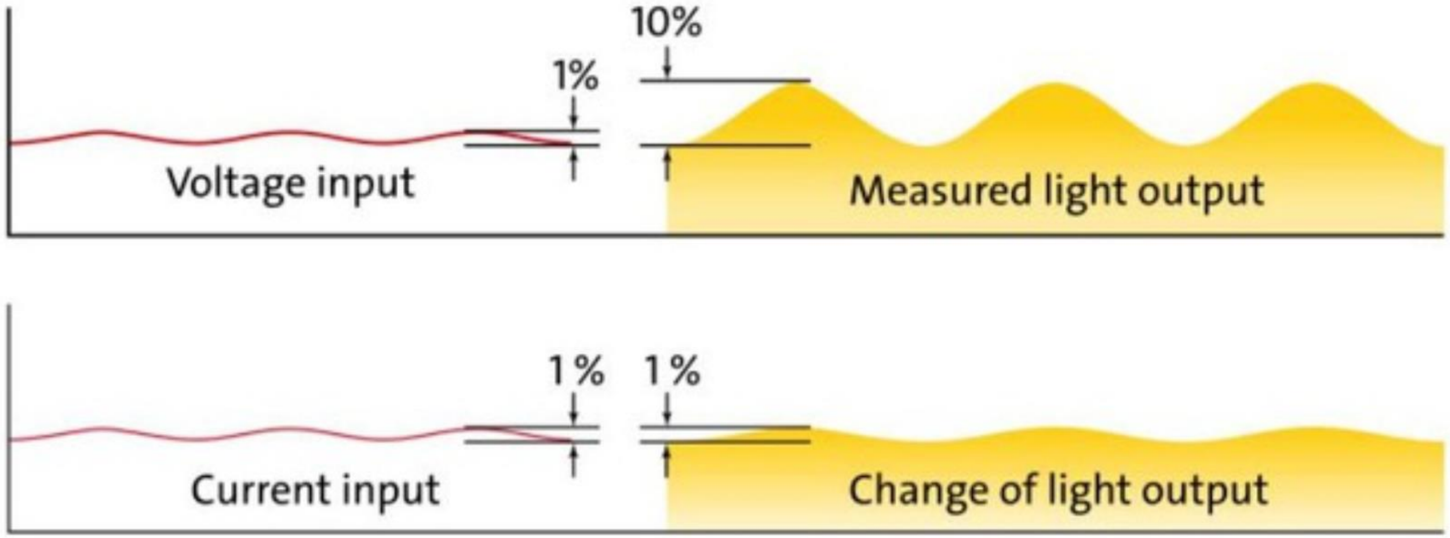
Strobing


Trigger pulse
Trigger pulse
Camera exposure
Camera exposure
LED strobe output
LED strobe output
Time

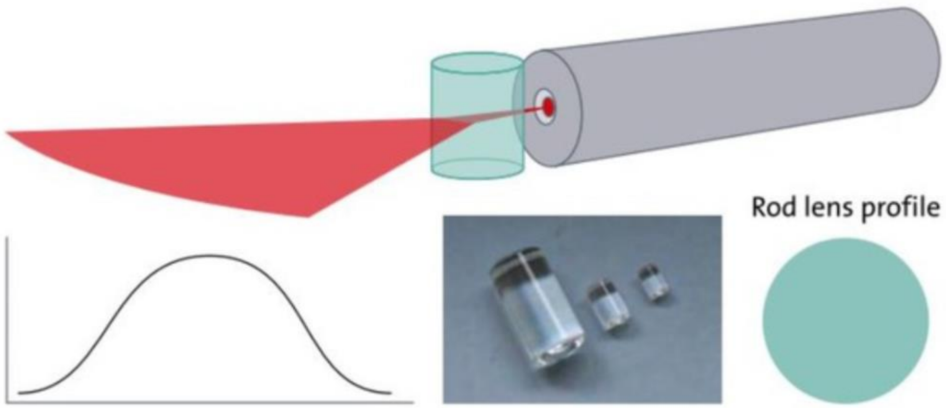
Overpowering

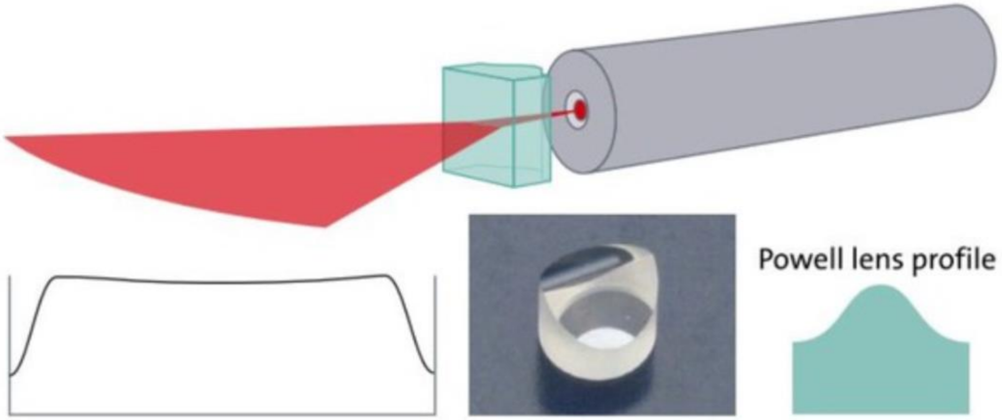


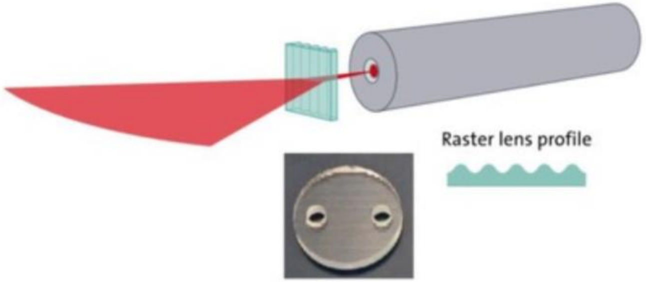
LED controller



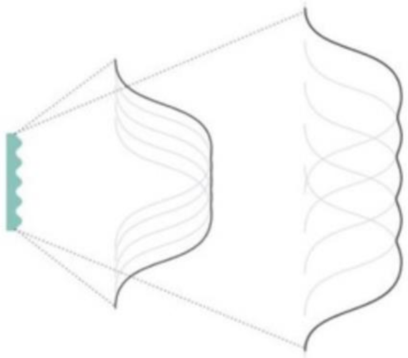
Laser lighting



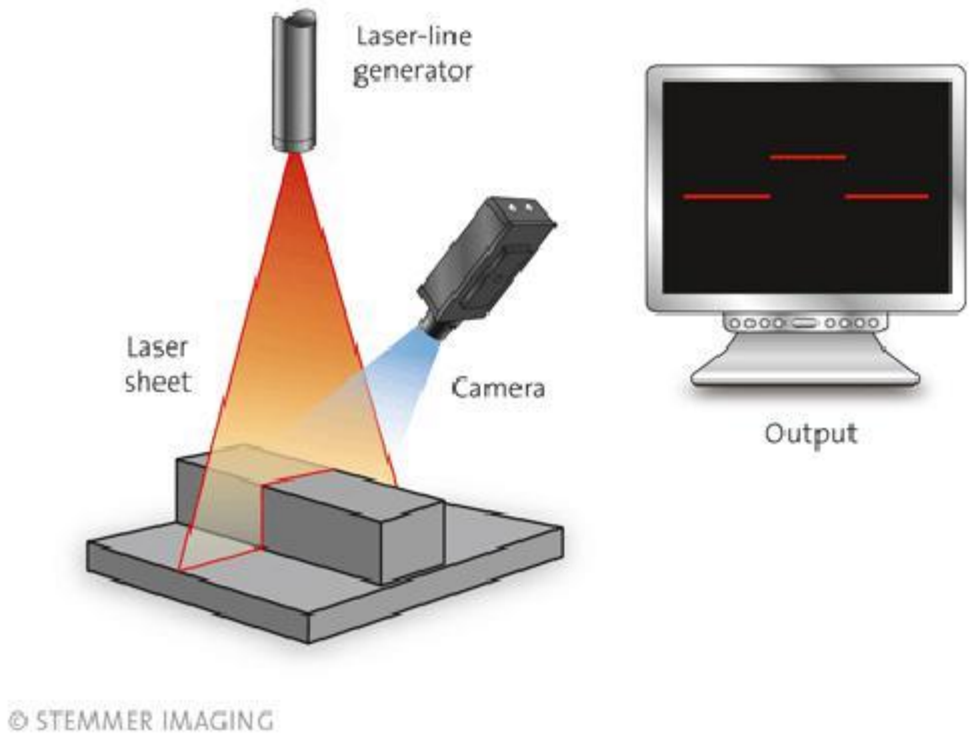




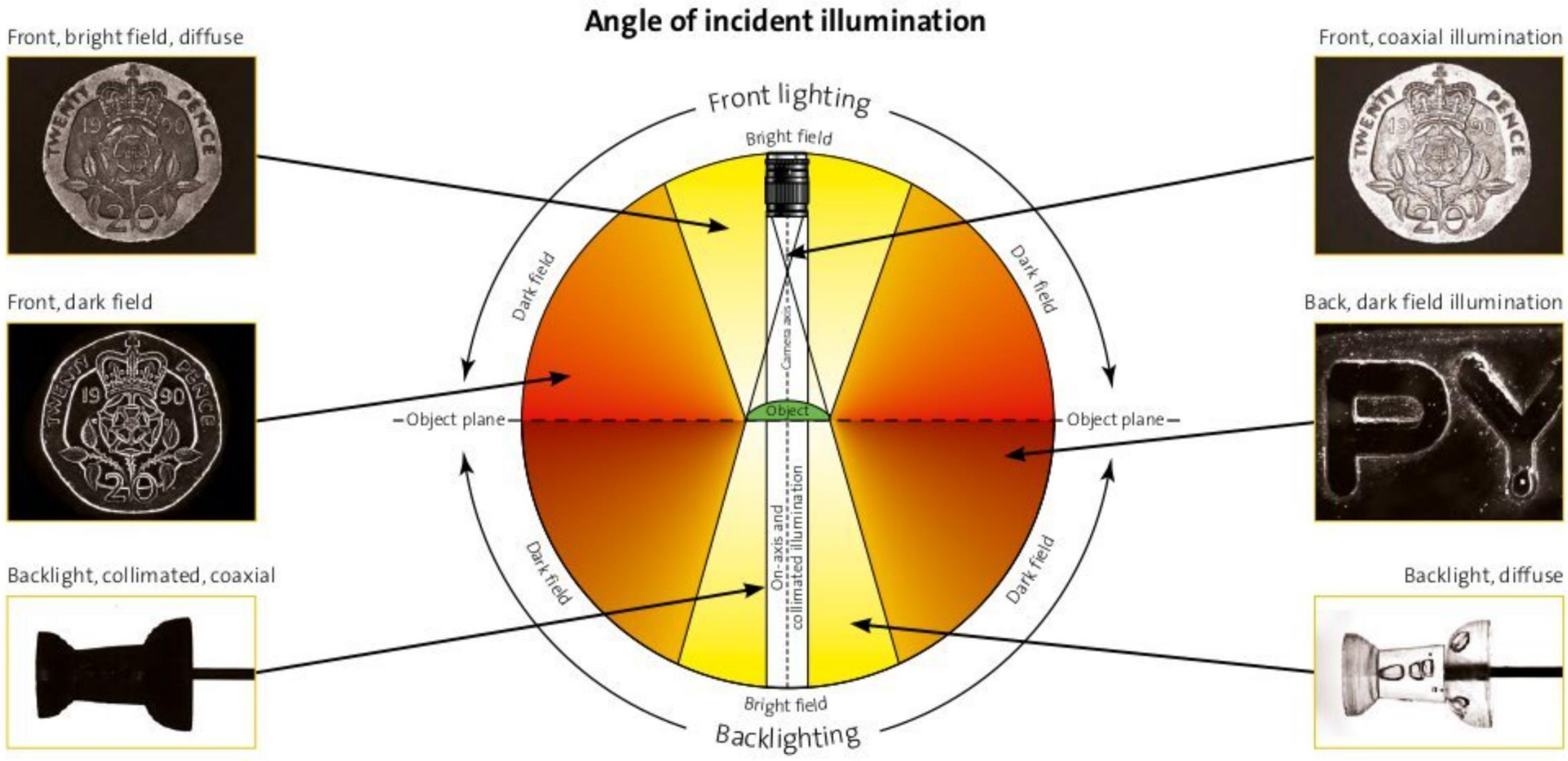
Raster lens profile
'pseudo-Gaussian' profile



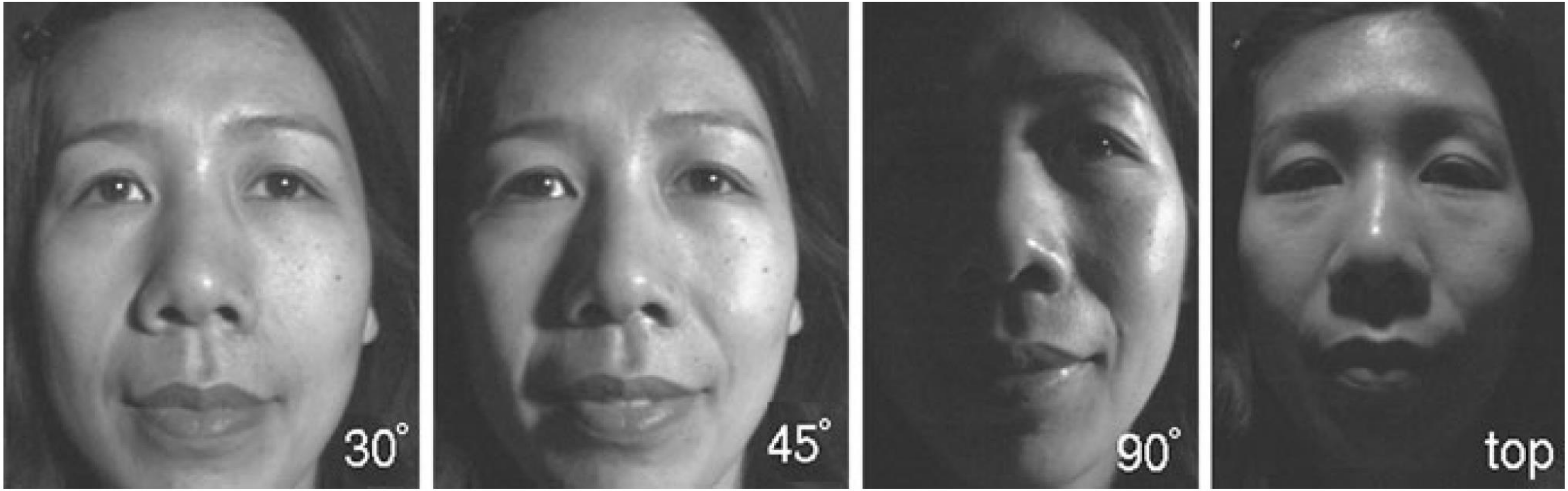
Application



Lighting overview

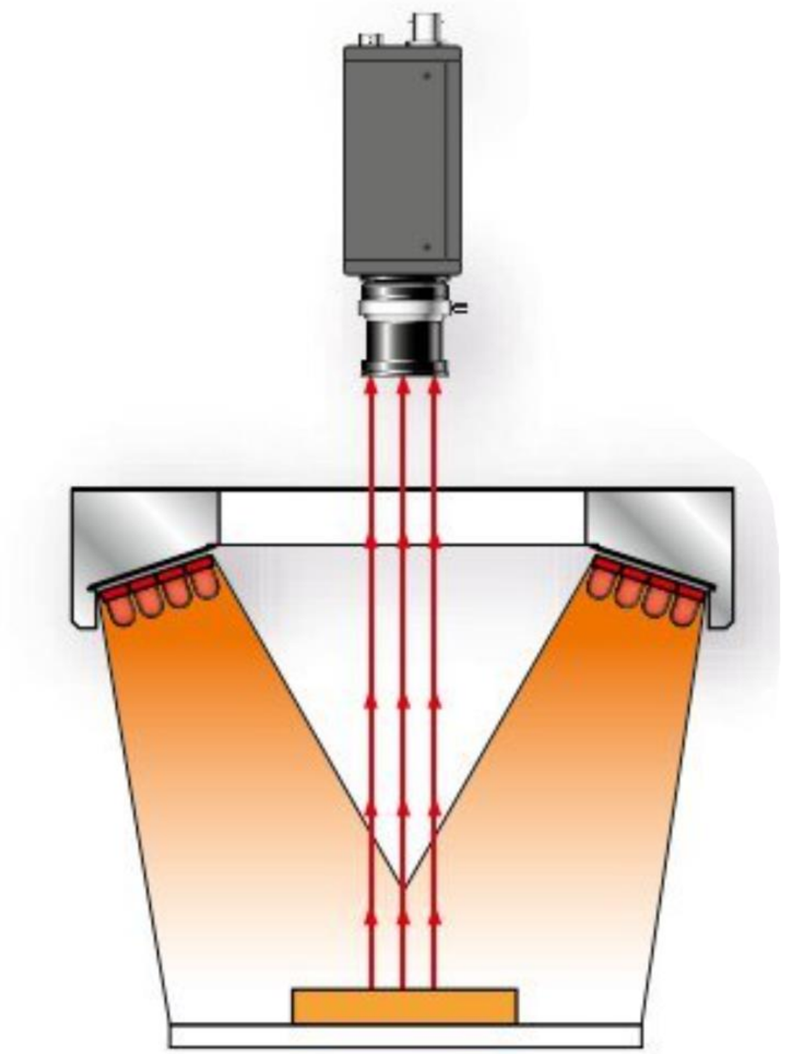


Illuminating from different directions

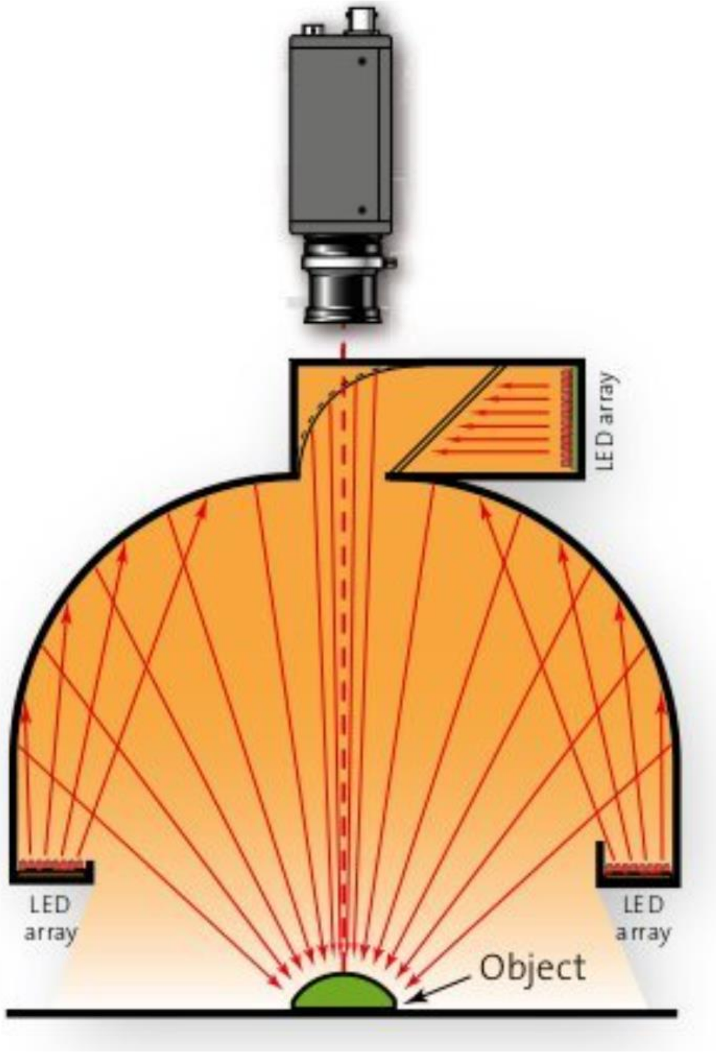


Backlighting

Ring lighting

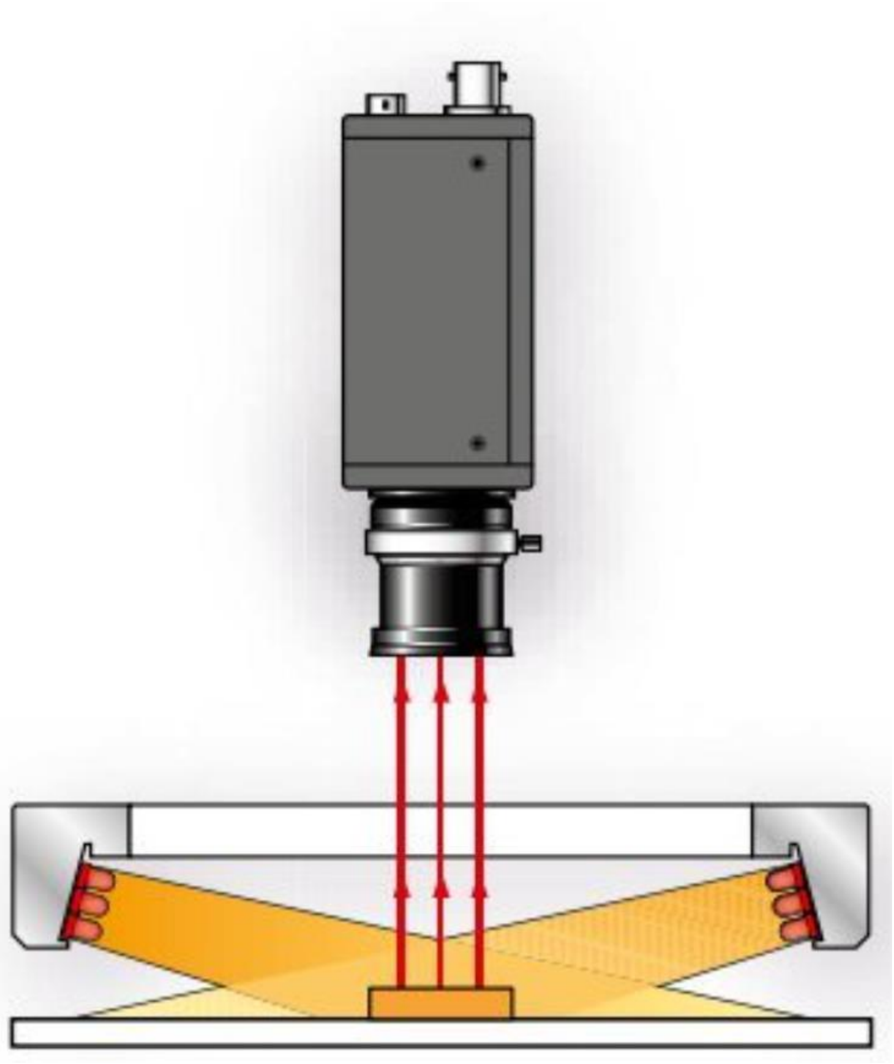


Dome lighting

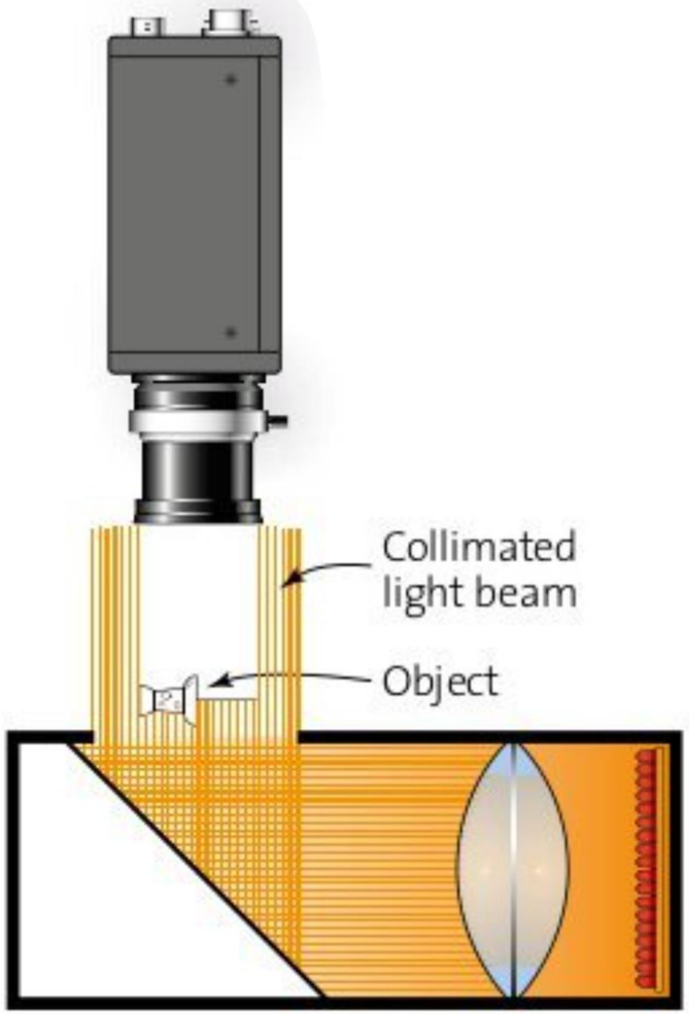


Ring vs dome lighting

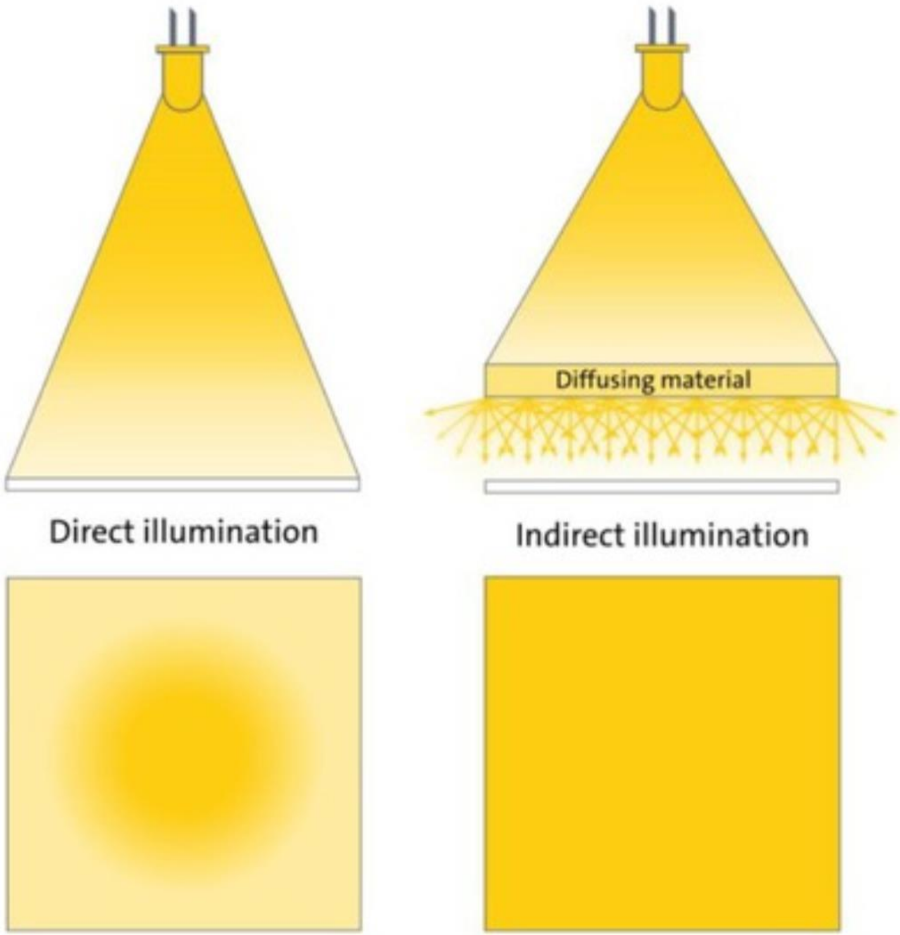
Dark field illumination



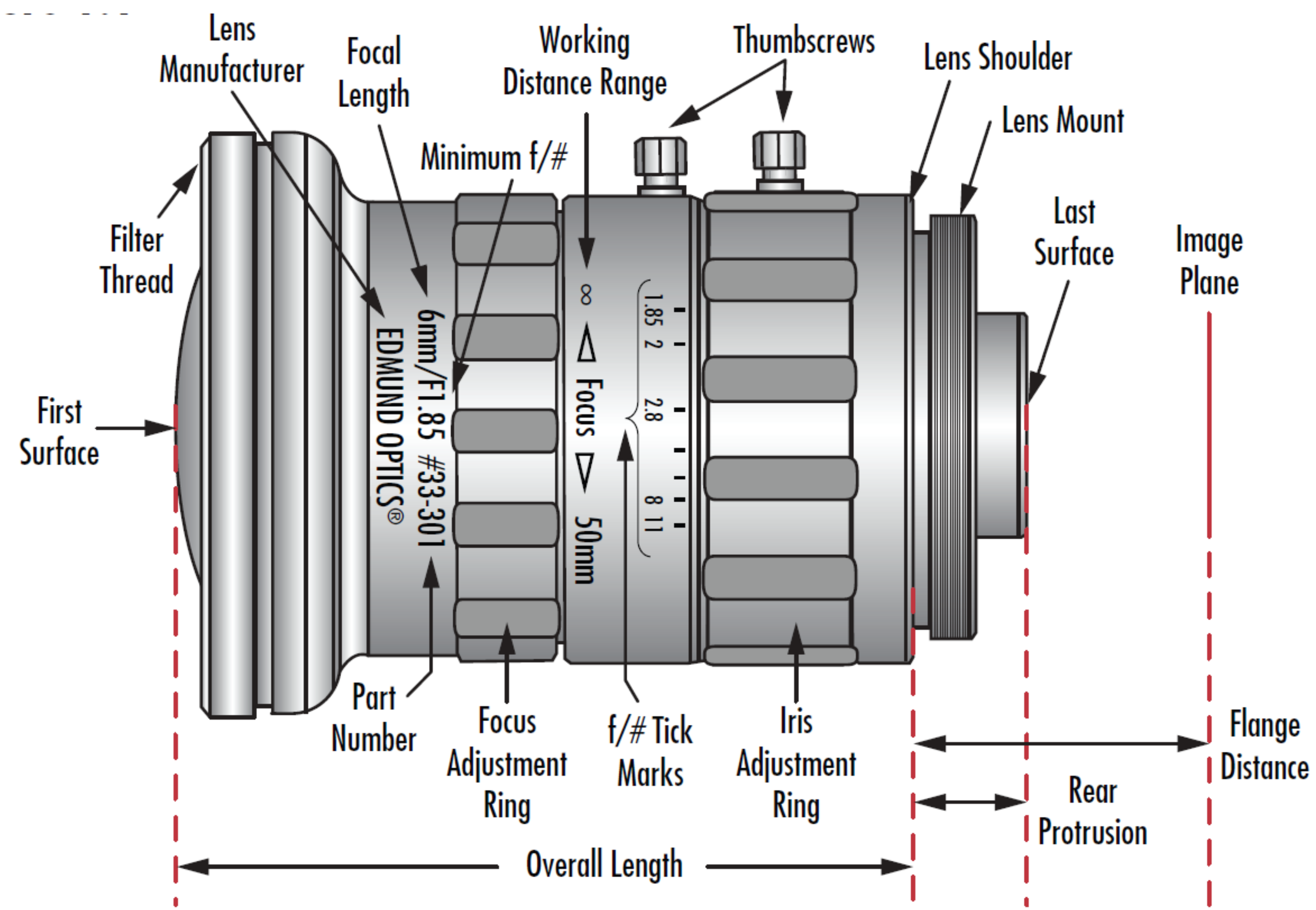
Back lighting



Direct illumination vs diffused light

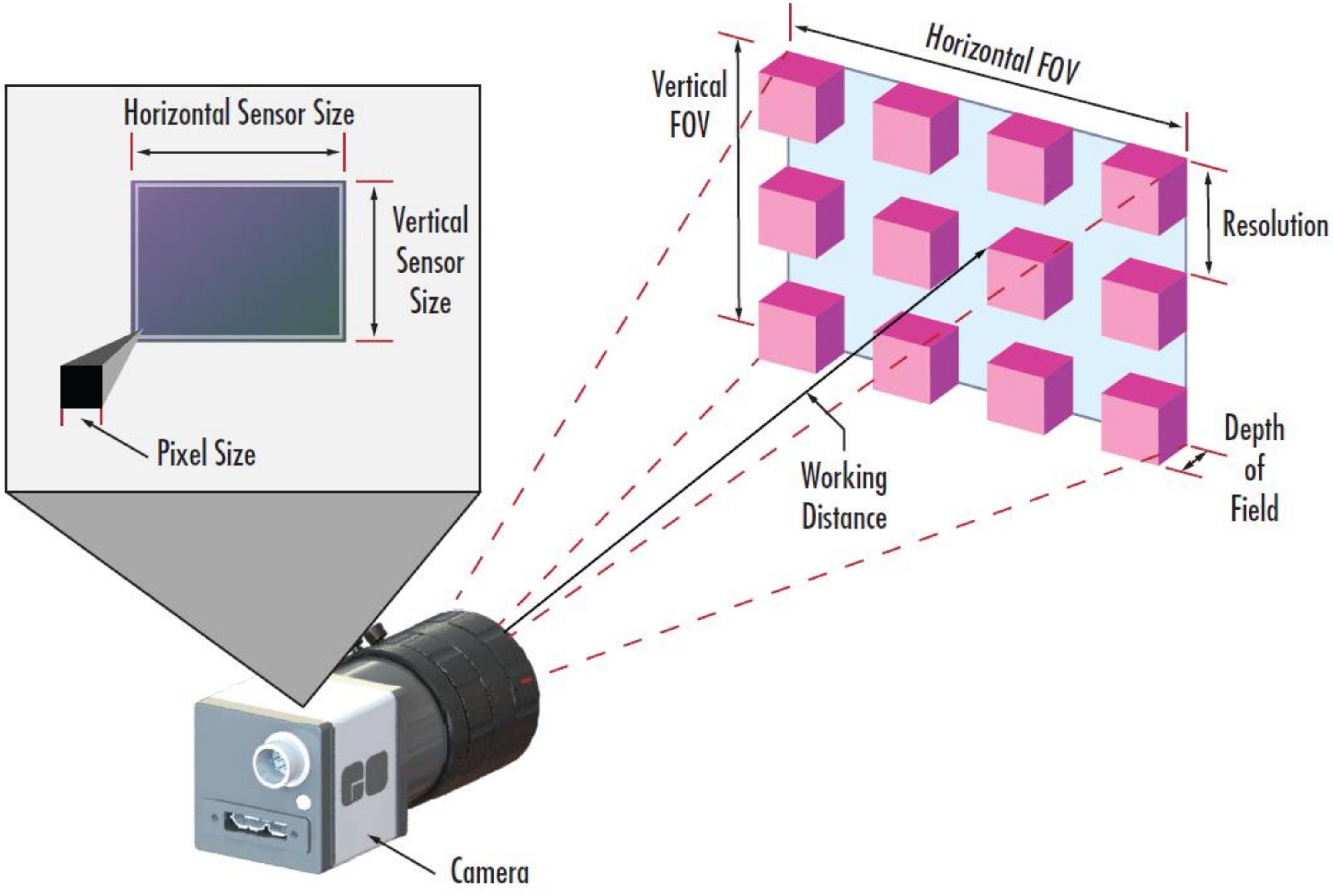


Lens



Lens Manufacturer
Focal Length
Working Distance Range
Thumbscrews
Lens Shoulder
Lens Mount
Minimum f/#
Filter Thread
Last Surface
Image Plane
First Surface
EDMUND OPTICS®
6mm/F1.85 #33-301
∞
Focus
50mm
1.85 2
2.8
8 11
Part Number
Focus Adjustment Ring
f/# Tick Marks
Iris Adjustment Ring
Flange Distance
Rear Protrusion
Overall Length

Fundamental Imaging Parameters 1



Fundamental Imaging Parameters 2

- **Field of View (FOV):** The viewable area of the object under inspection. This is the portion of the object that fills the camera's sensor. This area is commonly reduced to the horizontal (HFOV) or vertical (VFOV) dimension for ease of calculation.
- **Working Distance (WD):** The distance from the front or first surface of the lens to the object under inspection.
- **Resolution:** The minimum feature size of the object that can be distinguished by the imaging system typically specified as a spatial frequency in units of line pairs per millimeter (lp/mm).

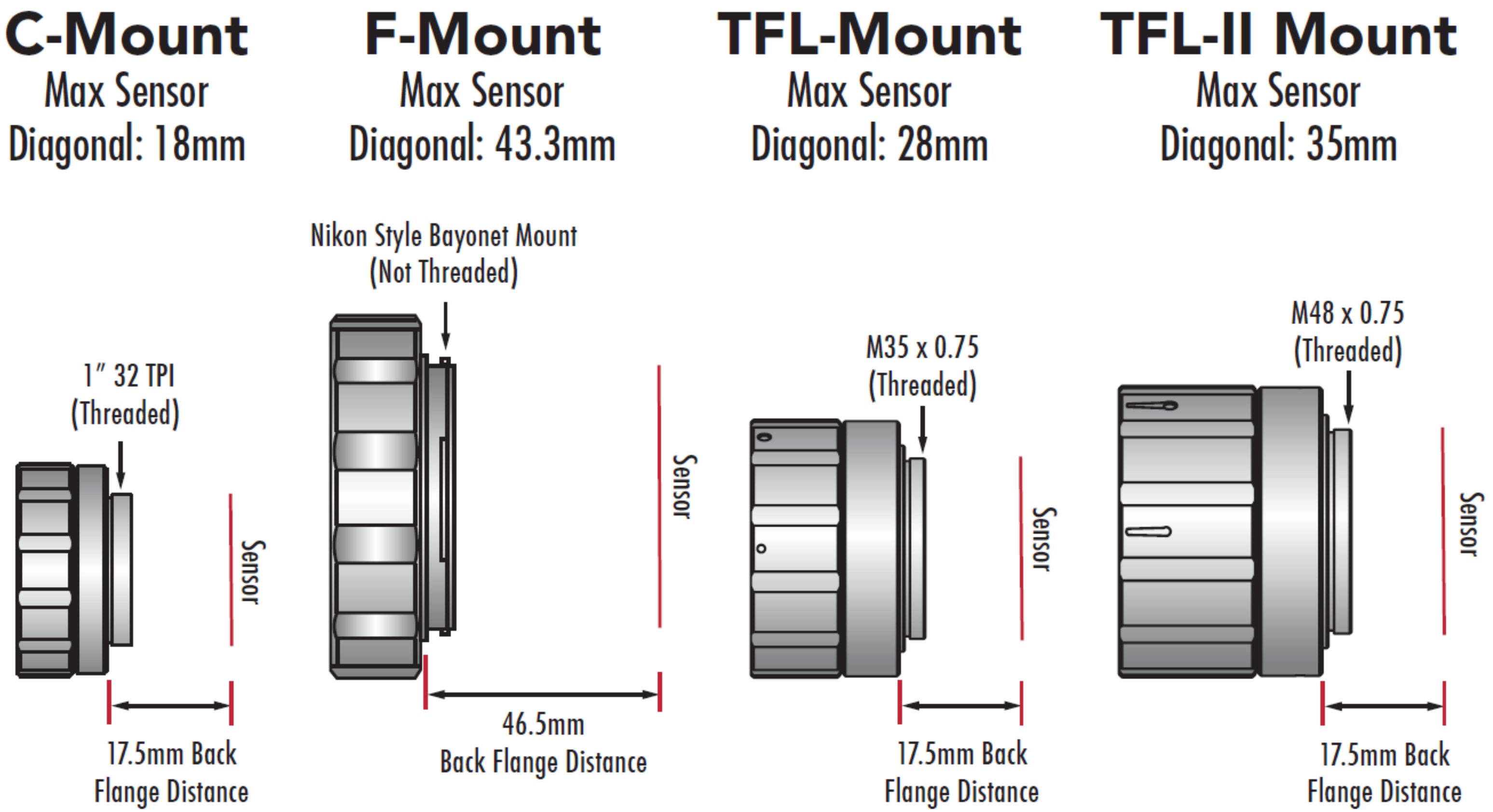
Fundamental Imaging Parameters 3

- **Depth of Field (DOF):** The maximum object depth that can be maintained entirely in acceptable focus. DOF is also the amount of object movement (in and out of best focus) allowable while maintaining focus. This is not to be confused with depth of focus, which refers to the location at which an image can be in focus on a sensor.
- **Sensor Size (H):** The size of a camera sensor’s active area, typically specified in the horizontal or vertical dimension. This parameter is important in determining the proper lens magnification required to obtain a desired field of view.

Fundamental Imaging Parameters 4

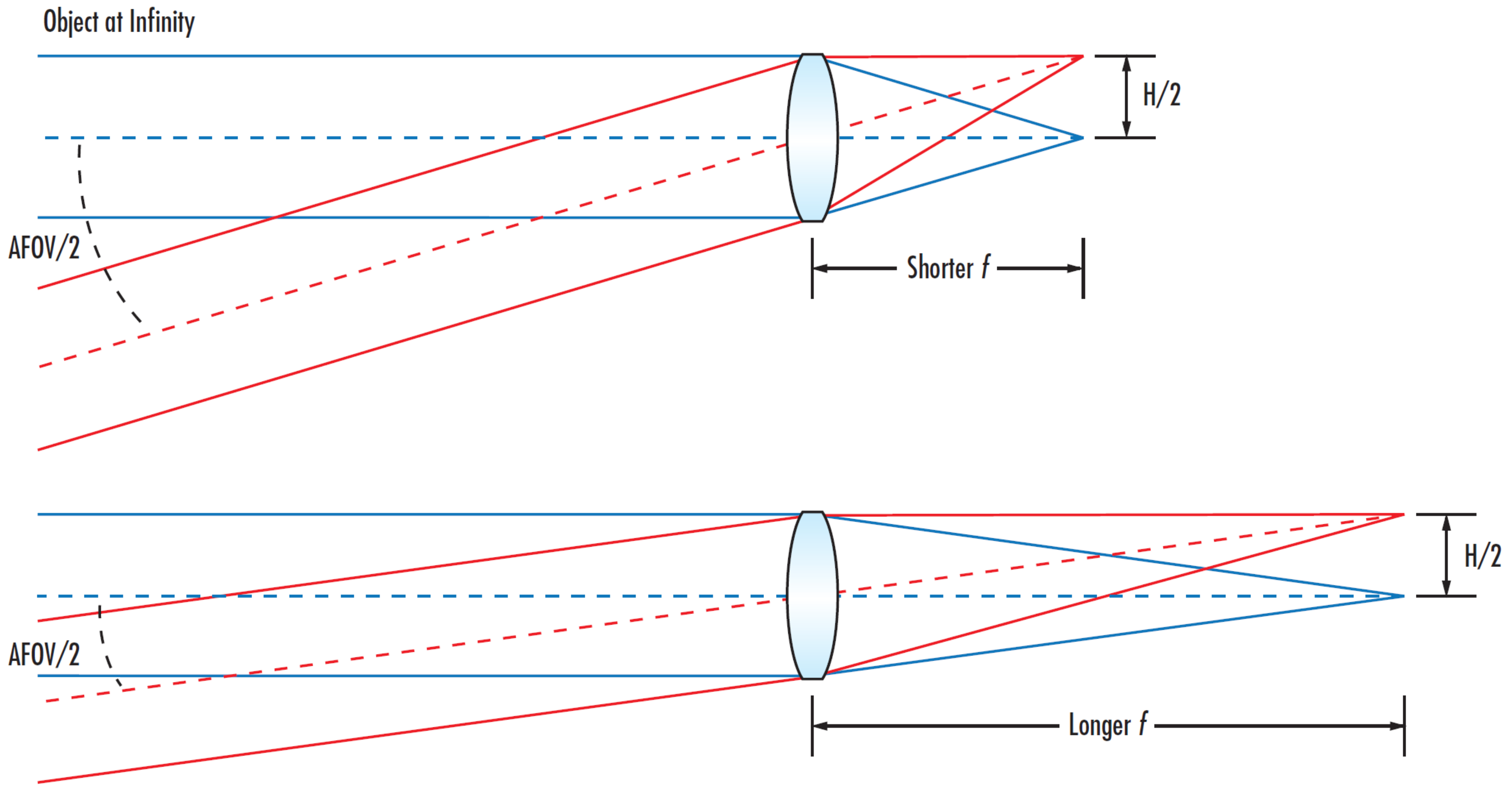
- **Pixel Size (*s*):** Typically square in shape, pixels make up the sensor grid and have a dimension in the order of micrometers ormicrons (µm).
- **Magnification (*m*):** The magnification of a lens, sometimes referred to as PMAG (**P**rimary **Mag**nification), is defi ned as the ratio between the sensor size and the FOV.

Lens mounts


C-Mount
Max Sensor
Diagonal: 18mm
F-Mount
Max Sensor
Diagonal: 43.3mm
TFL-Mount
Max Sensor
Diagonal: 28mm
TFL-II Mount
Max Sensor
Diagonal: 35mm
1" 32 TPI
(Threaded)
Sensor
17.5mm Back
Flange Distance
Nikon Style Bayonet Mount
(Not Threaded)
Sensor
46.5mm
Back Flange Distance
M35 x 0.75
(Threaded)
Sensor
17.5mm Back
Flange Distance
M48 x 0.75
(Threaded)
Sensor
17.5mm Back
Flange Distance

Types of lenses

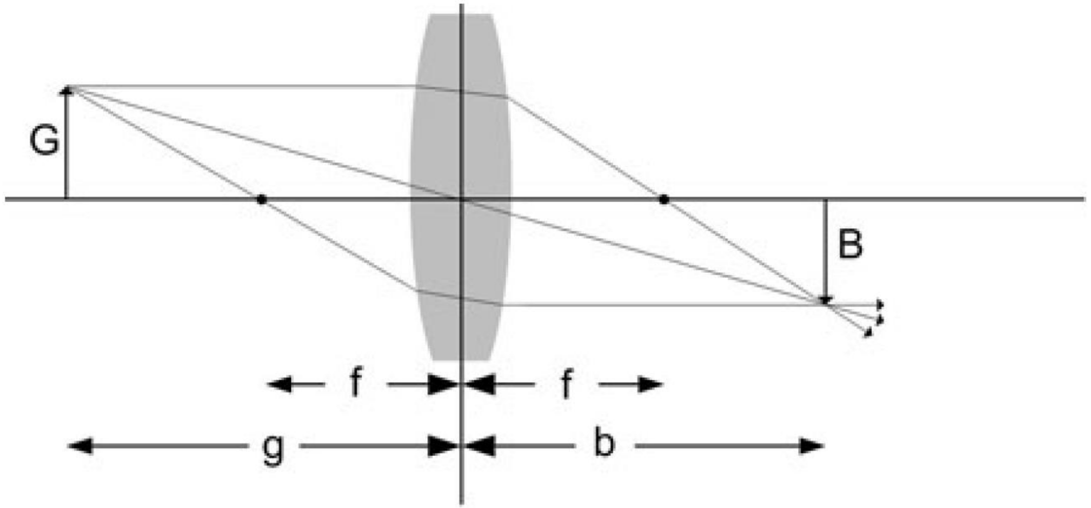
- A **fixed focal length lens**, also known as a conventional or entocentric lens, is a lens with a fixed **angular field of view** (AFOV).
- A **varifocal lens** or a **zoom lens**; these types of lenses allow for adjustment of their focal lengths and thus have variable AFOV.

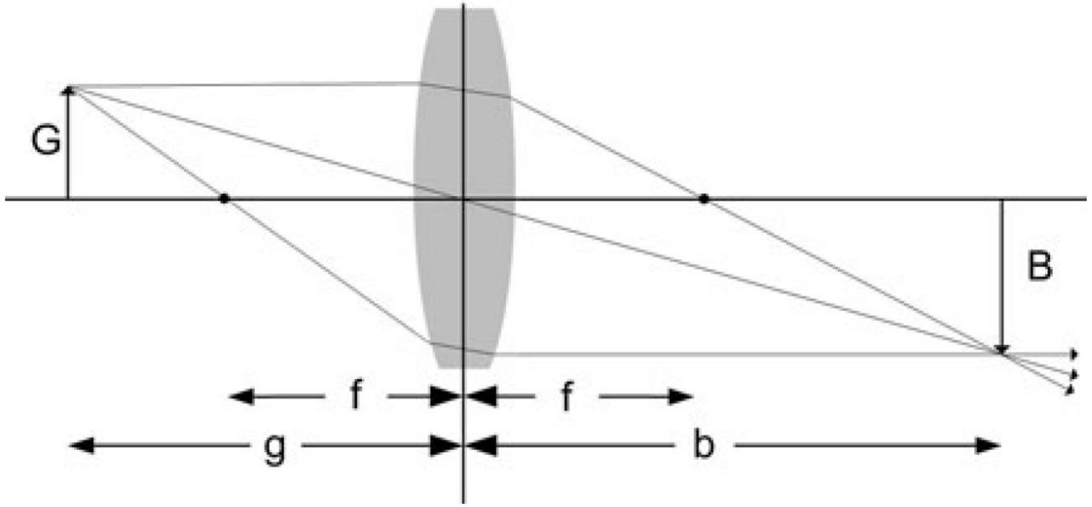


Equations

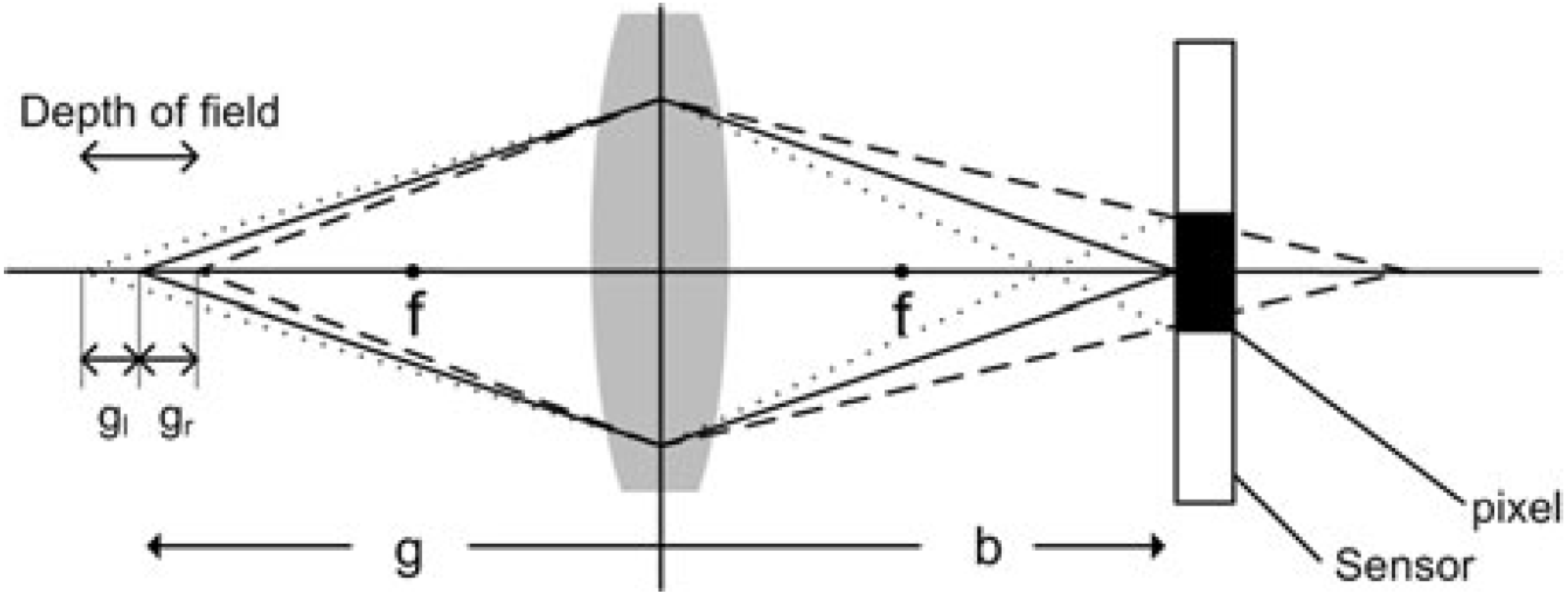
- $AFOV = 2 \cdot tan^{-1}\left(\frac{H}{2f}\right)$
- $AFOV = 2 \cdot tan^{-1}\left(\frac{FOV}{2WD}\right)$
- $\frac{1}{p} + \frac{1}{q} = \frac{1}{f}$
- $\frac{1}{f} = (n-1)\left(\frac{1}{R_1} + \frac{1}{R_2}\right)$

Zoom – different focal distances

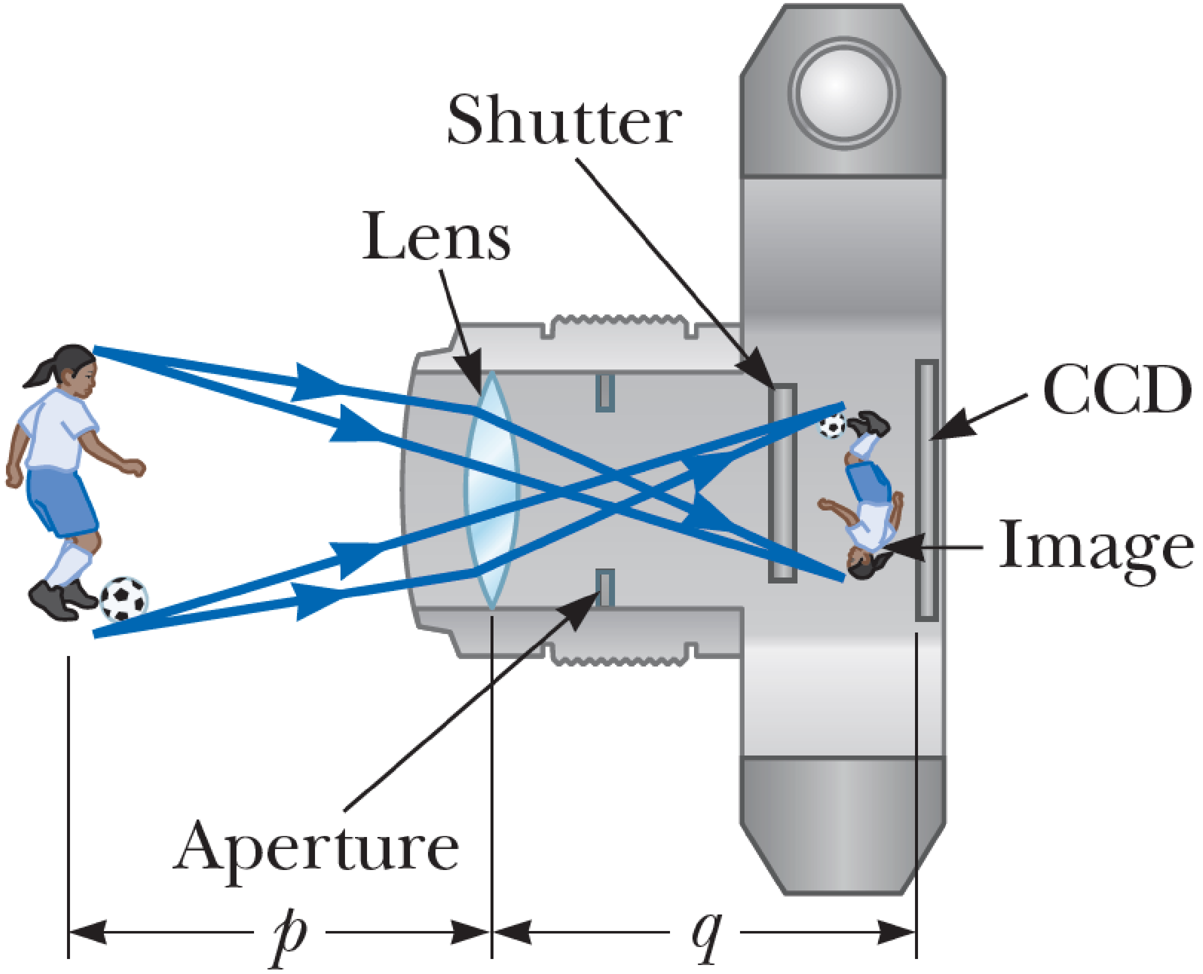




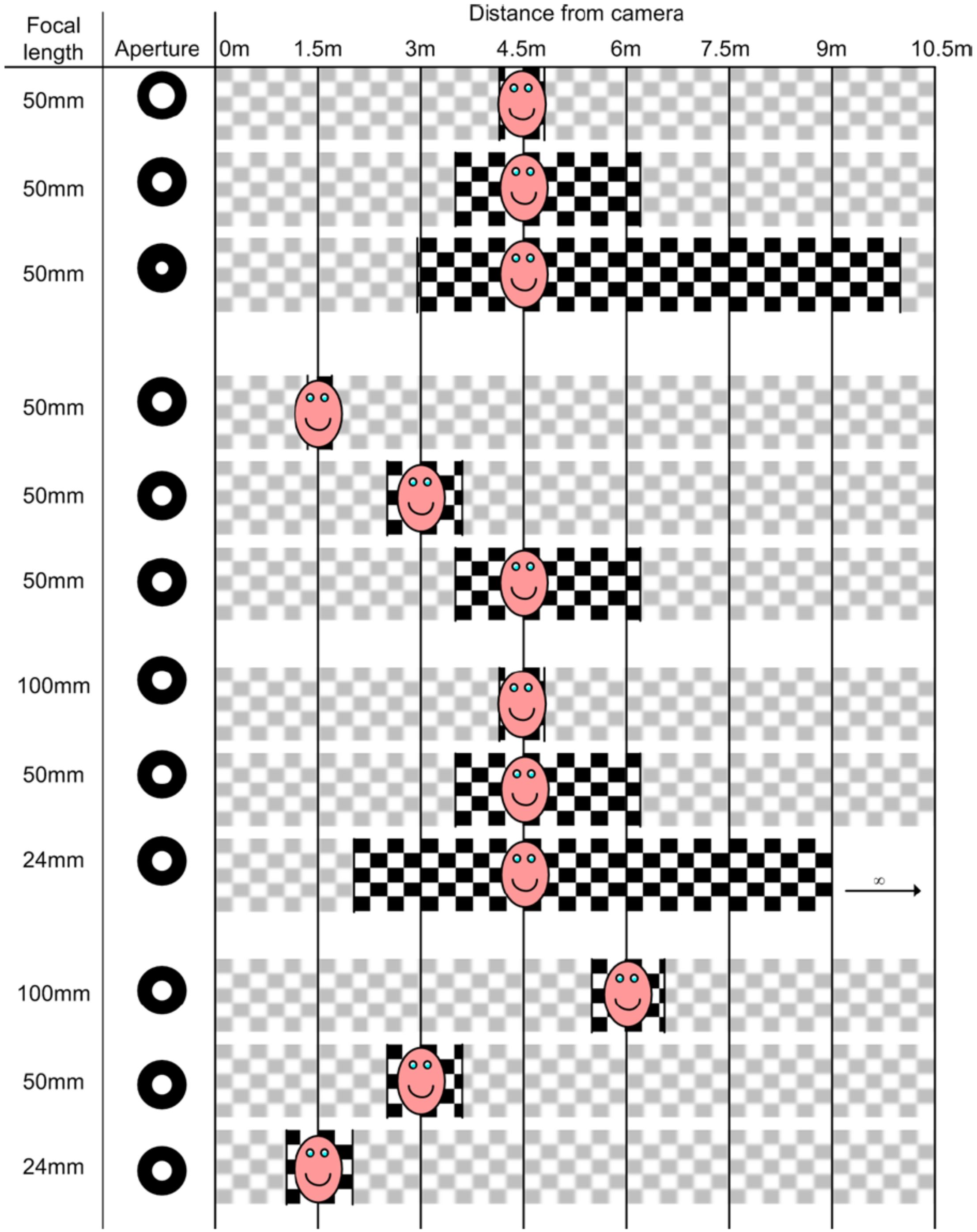
Depth of field



Camera



Comparison



Focal length
Aperture
Distance from camera
0m
1.5m
3m
4.5m
6m
7.5m
9m
10.5m
50mm
50mm
50mm
50mm
50mm
50mm
100mm
50mm
24mm
∞
100mm
50mm
24mm

f-number

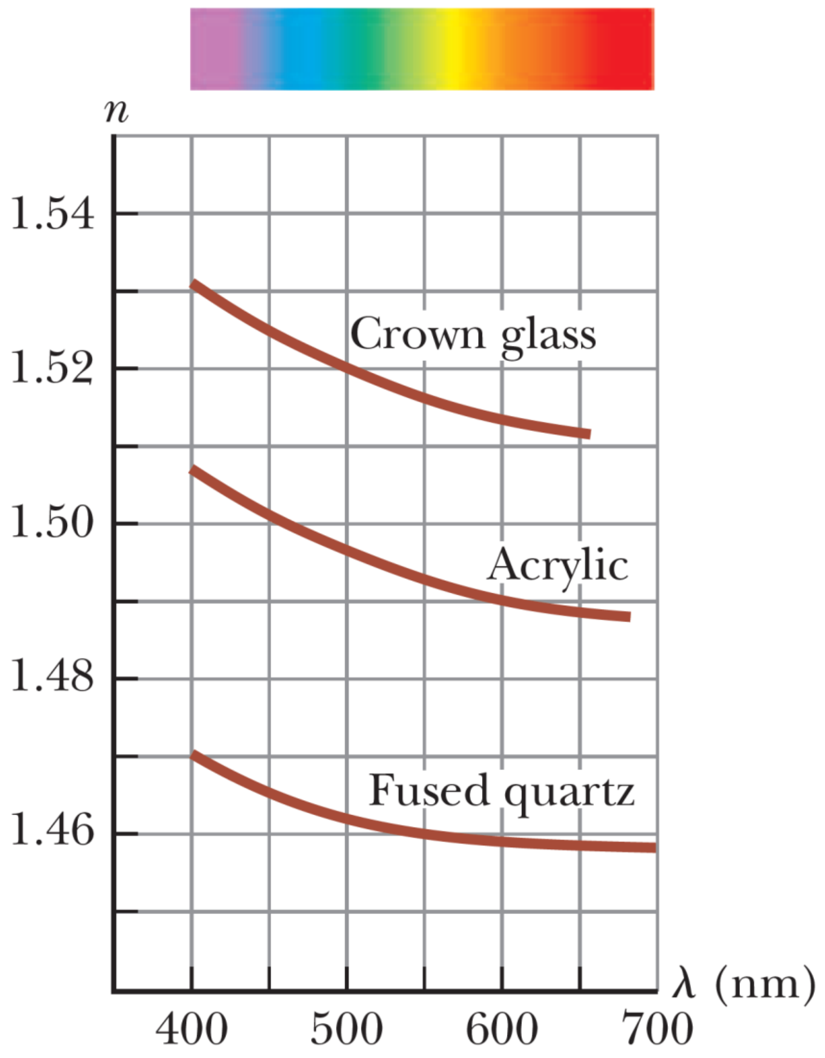
- f-number=f/D
- Intensity (I)

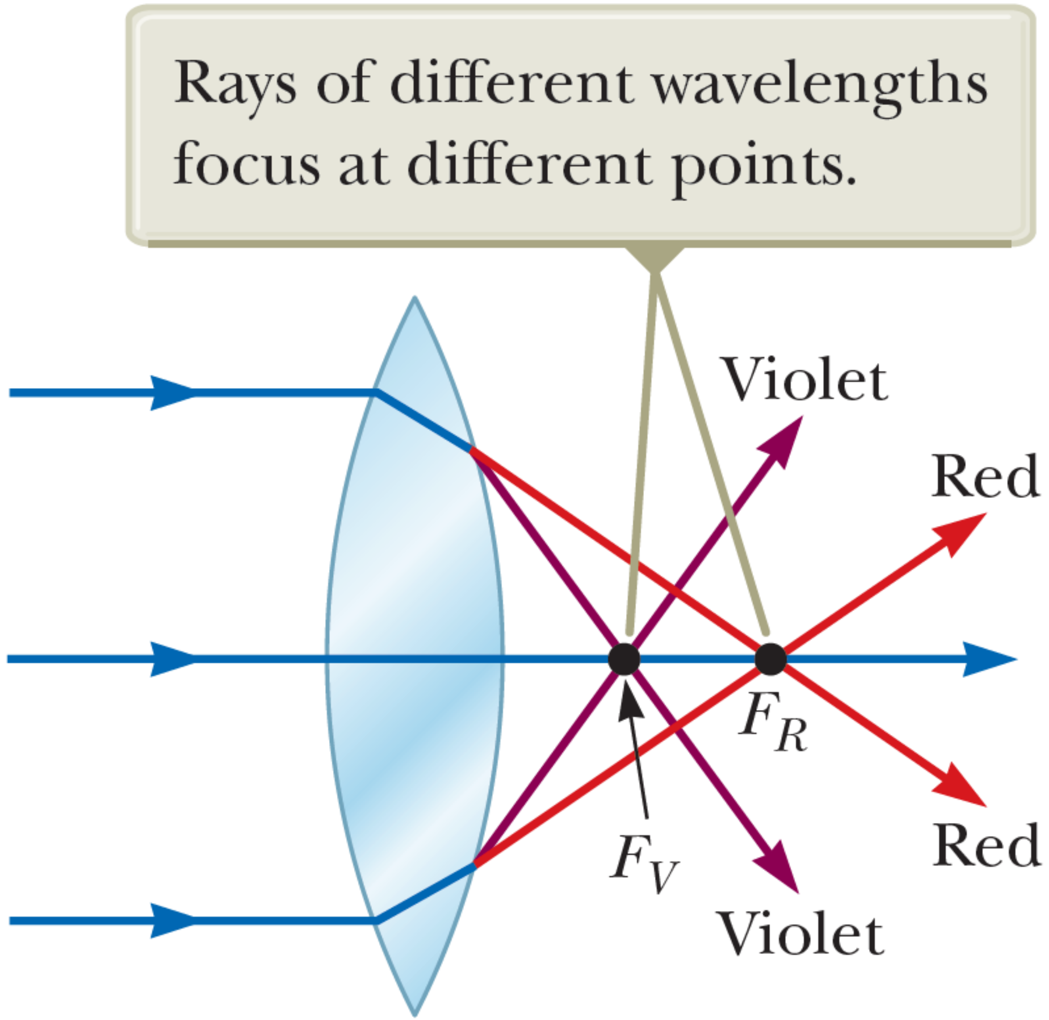
$$I \propto 1/\left(\frac{f}{D}\right)^2 \propto \frac{1}{f\!-\!number^2}$$

Best f-number lens

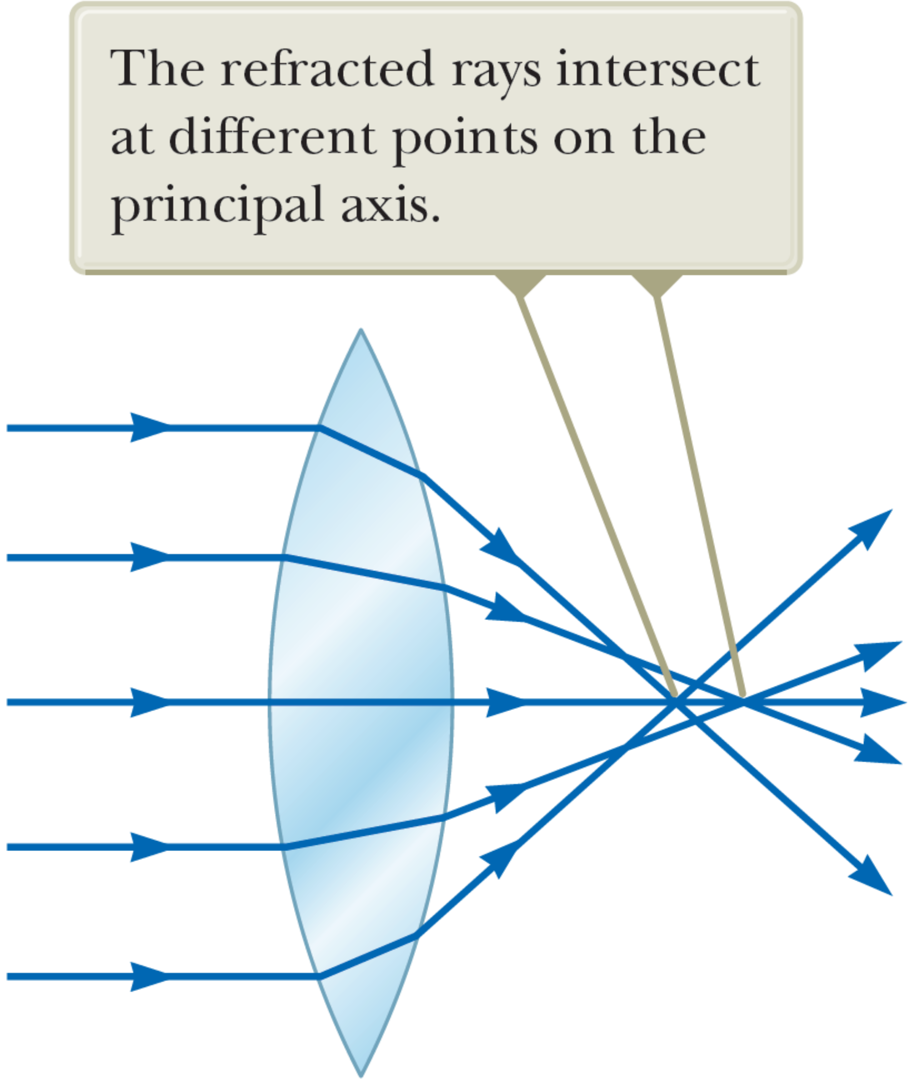
- Zeiss f/0.7 (build for Nasa)
- Leica Noctilux f/0.95 (price cca. 10.000EUR)

Chromatic aberration

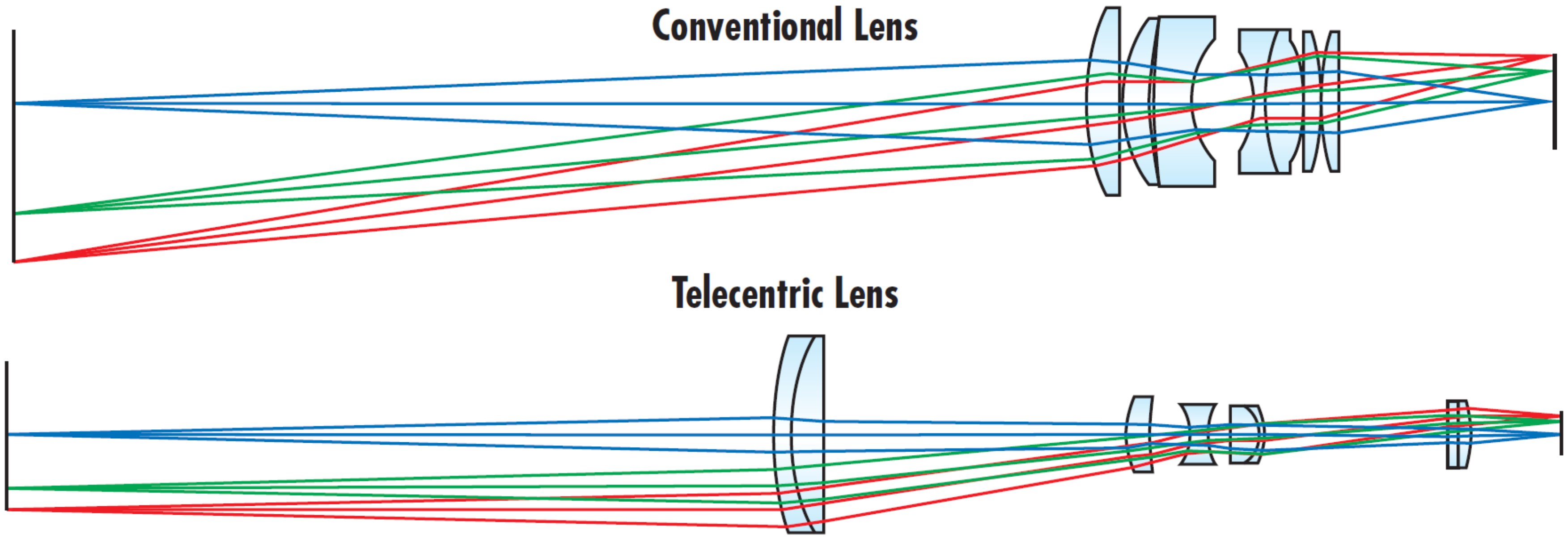


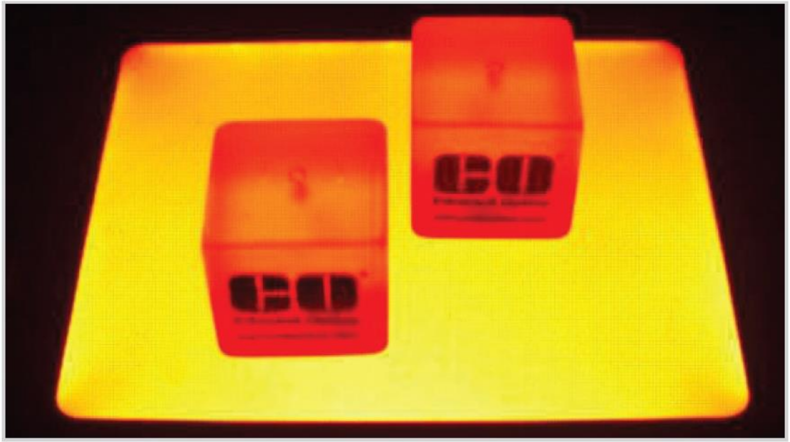


Spherical aberration

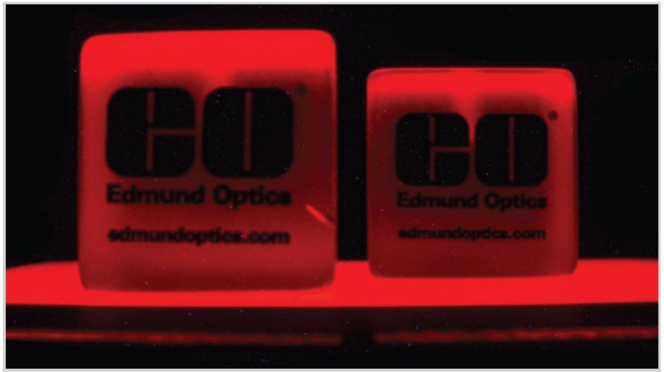


Telecentric lens



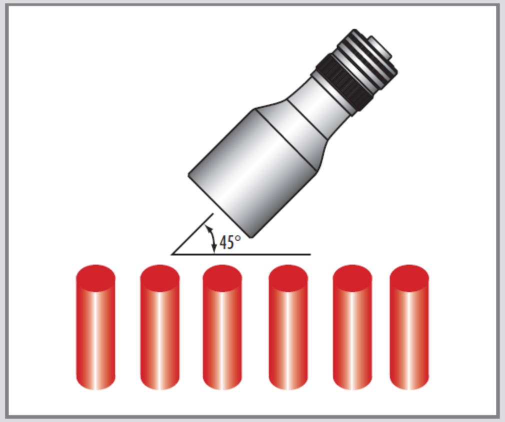


Fixed Focal Length Lens

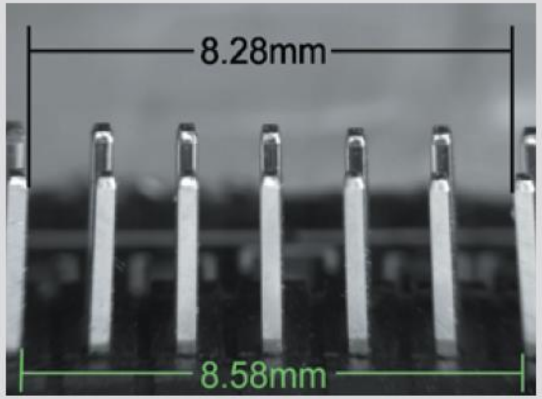


Telecentric Lens

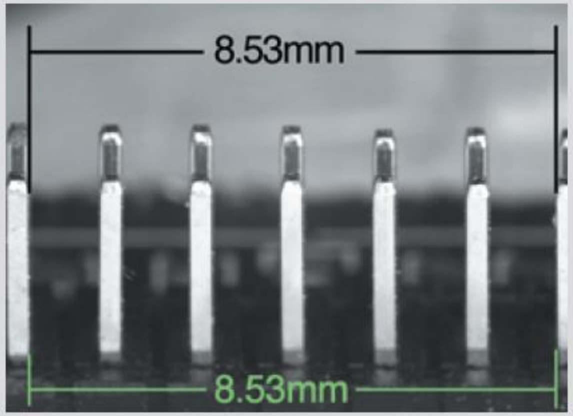
Distortion



Arrangement



Fixed Focal Length Lens



Telecentric Lens

Influence of exposure time

Influence of aperture (f-number)

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

Image formation

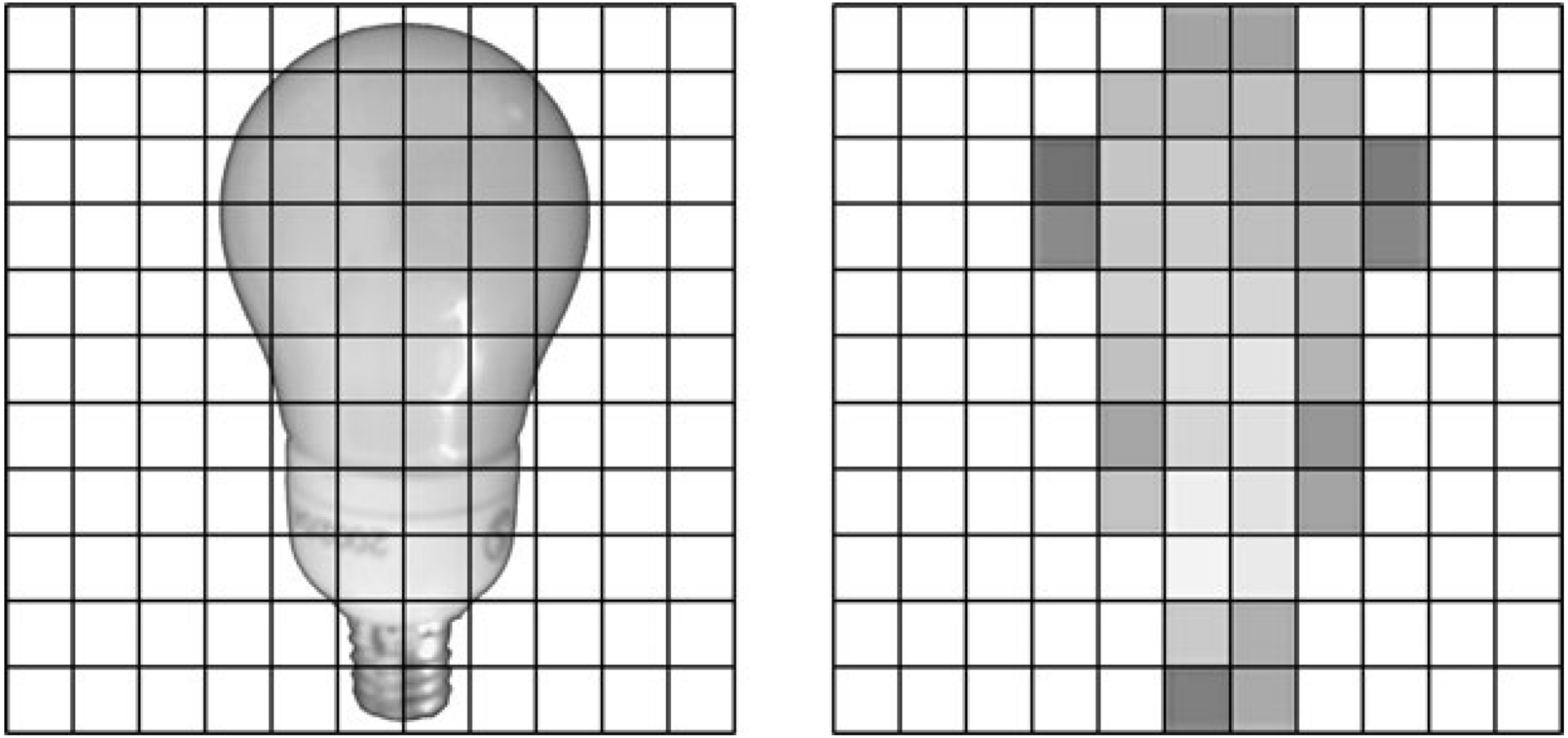
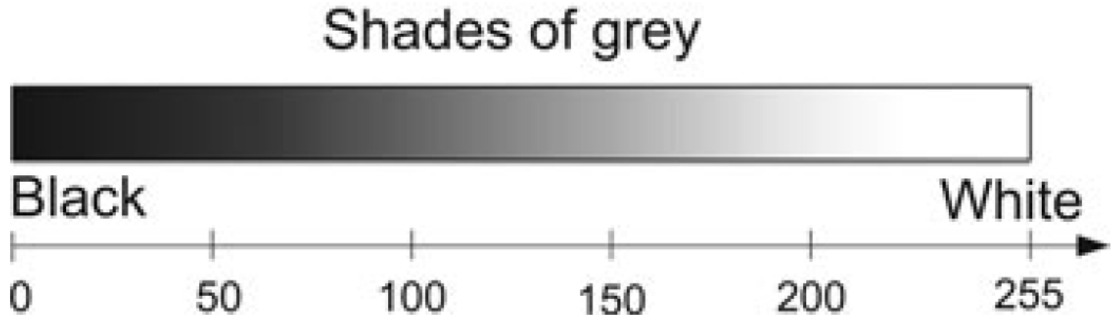
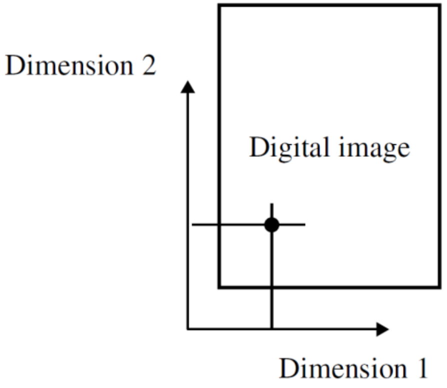


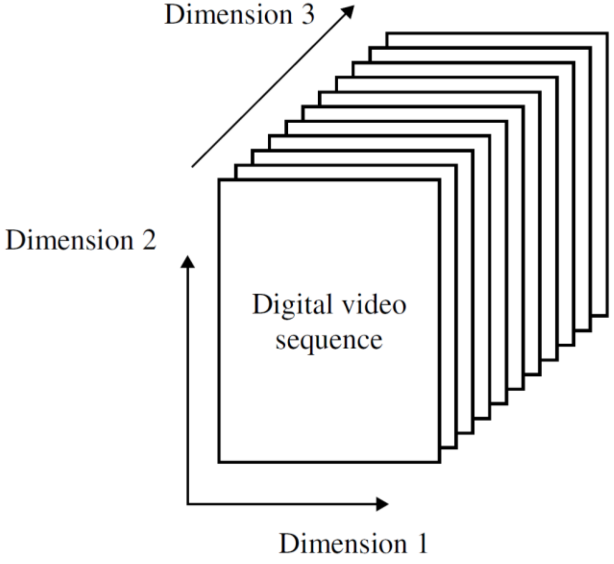
Image in “mathematical sense”



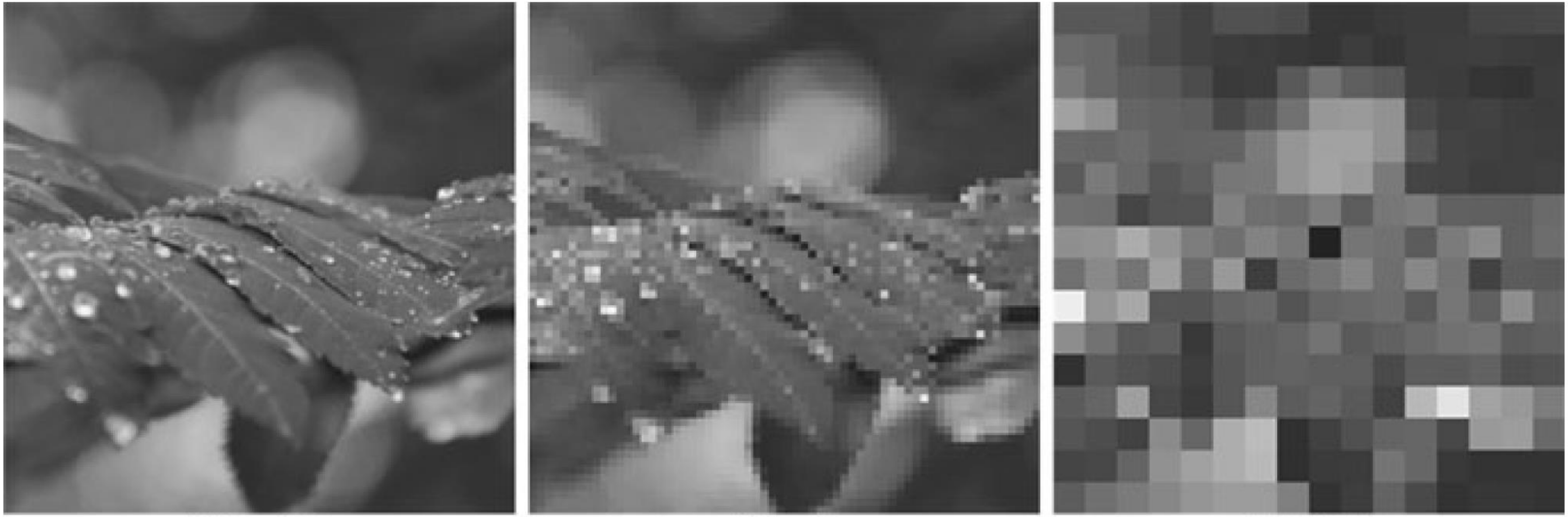
y \ x	0	1	2	3
0	10	20	12	23
1	17	100	25	95
2	9	17	8	22
3	16	89	19	92

The dimensionality of images and video



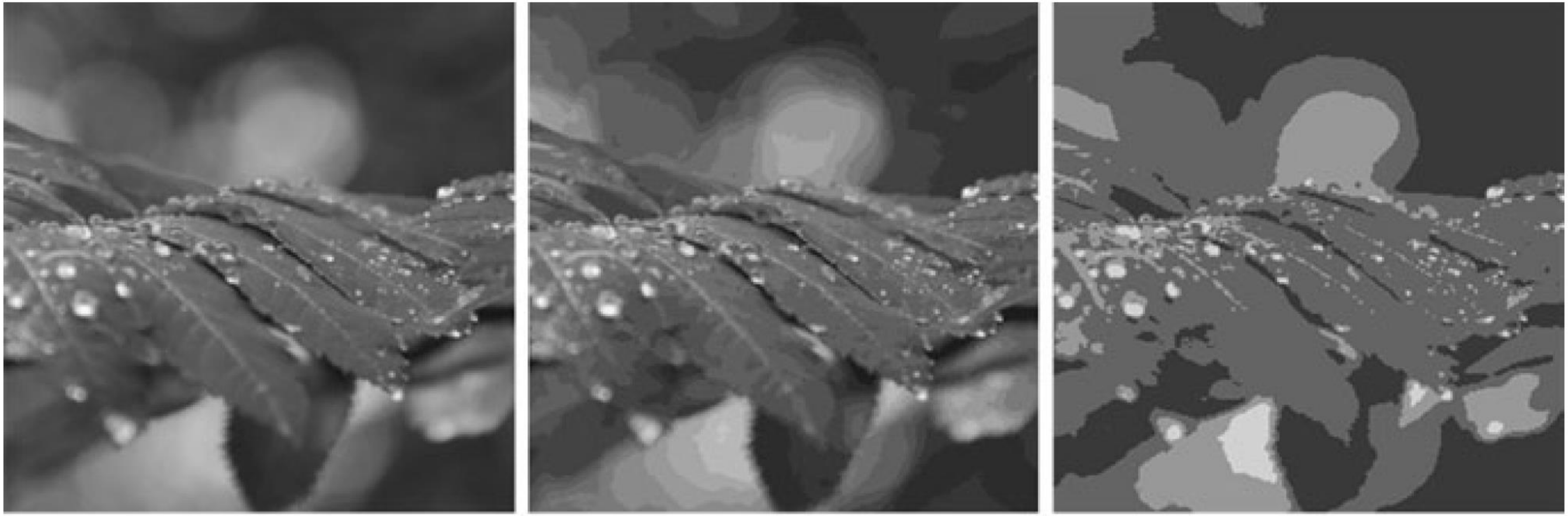


Spatial resolution (256x256, 64x64, 16x16)



256 x 256 | 64 x 64 | 16 x 16

Gray level resolution



256 gray-levels | 16 gray-levels | 4 gray-levels

Image size

- 256 x 256 x 8bit (depth) = 2^8 x 2^8 x 1 byte = 2^{16} bytes = 65 kB

Comparison (spatial resolution)

Figure 1.4	(a)		(b)		(c)		(d)		(e)		(f)
Spatial resolution	512×512		256×256		128×128		64×64		32×32		16×16
Number of gray levels	256		256		256		256		256		256
Amount of data/bit	67108864		16777216		4194304		1048576		262144		65536
Ratio of data volume		4::1		4::1		4::1		4::1		4::1	

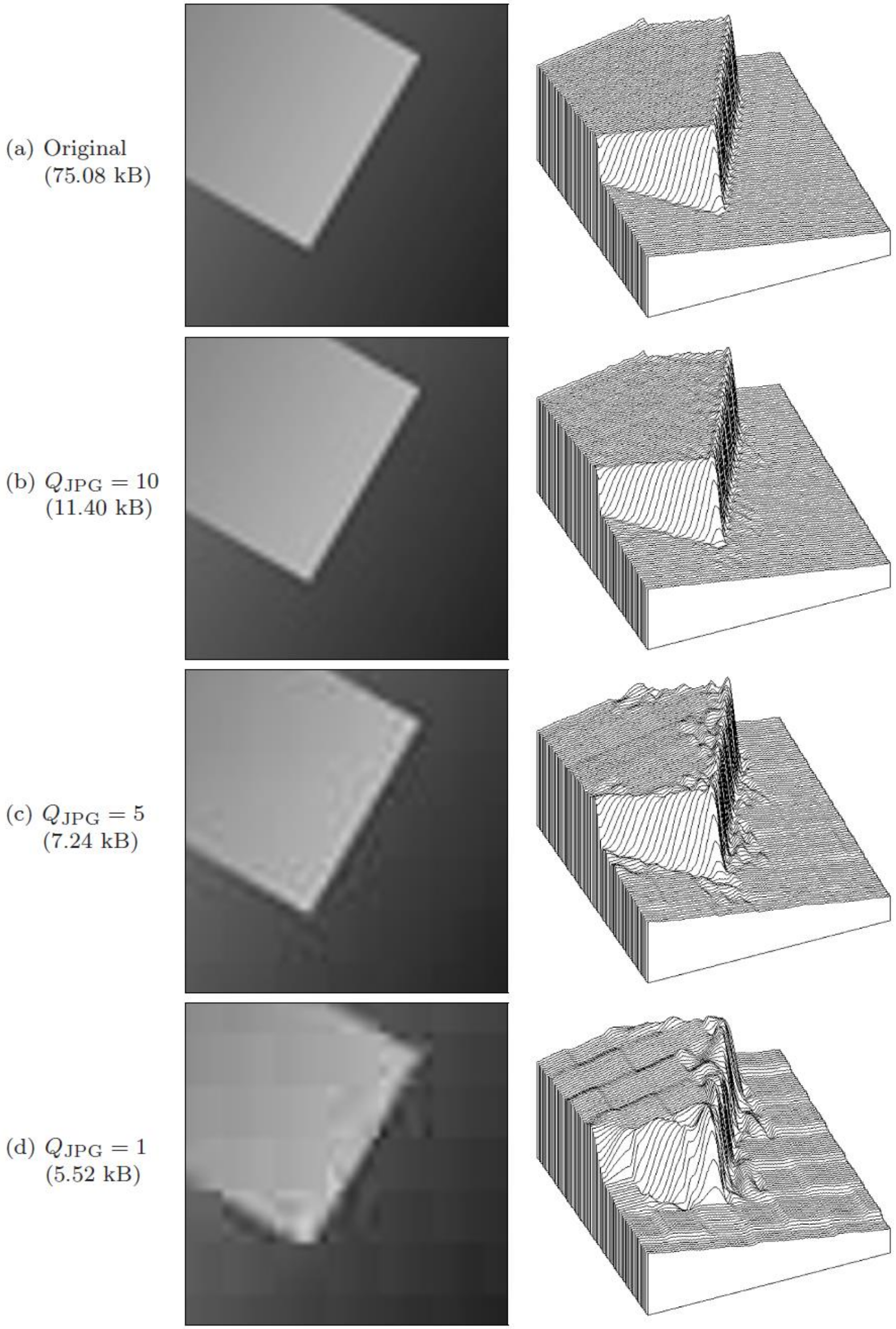
Comparison (image amplitude resolution; depth)

Figure 1.5	(a)		(b)		(c)		(d)		(e)		(f)
Spatial resolution	512×512		512×512		512×512		512×512		512×512		512×512
Number of gray levels	256		64		16		8		4		2
Amount of data/bit	67108864		16777216		4194304		2097152		1048576		524288
Ratio of data volume		4::1		4::1		2::1		2::1		2::1	

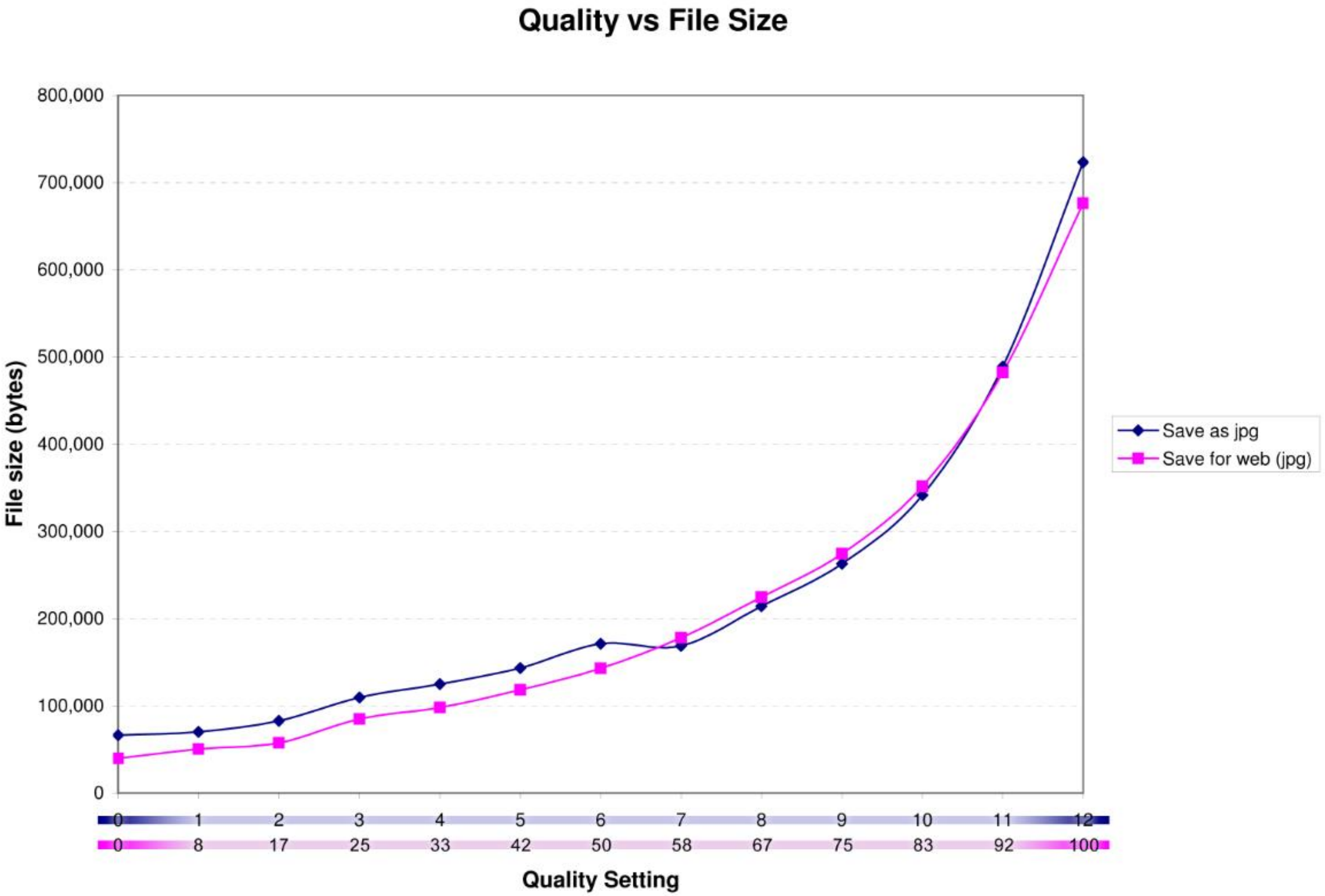
Simultaneous variation

Figure 1.6	(a)	(b)	(c)	(d)	(e)	(f)
Spatial resolution	256×256	181×181	128×128	90×90	64×64	45×45
Number of gray levels	256	64	32	16	8	4
Amount of data/bit	16777216	2096704	524288	129600	32768	8100
Ratio of data volume		8::1	4::1	4::1	4::1	4::1

Image compression

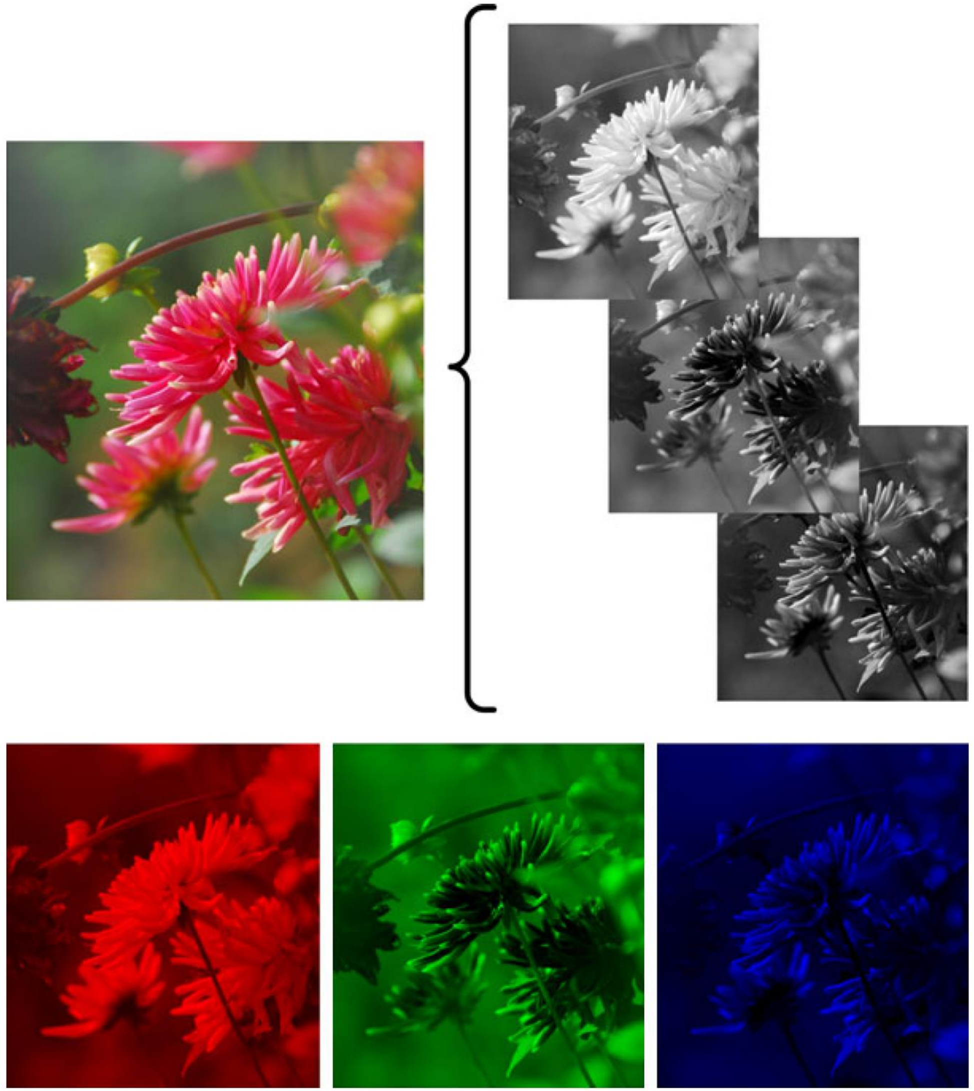


Comparison


Quality vs File Size
File size (bytes)
800,000
700,000
600,000
500,000
400,000
300,000
200,000
100,000
0
0 1 2 3 4 5 6 7 8 9 10 11 12
0 8 17 25 33 42 50 58 67 75 83 92 100
Quality Setting
Save as jpg
Save for web (jpg)


Saved as jpg
0 quality, 64.6KB
6 quality, 167.1KB
12 quality, 706.4KB
Saved for web jpg
0 quality, 38.8KB
50 quality, 139.5KB
100 quality, 660.4KB

Color images

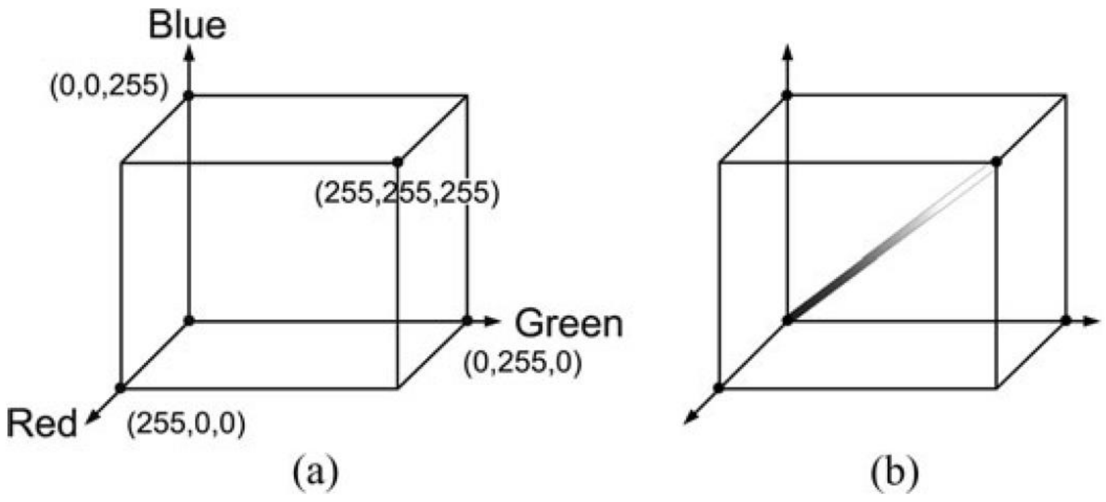


Why colors?

Photoreceptor cell	Wavelength in nanometers (nm)	Peak response in nanometer (nm)	Interpretation by the human brain
Cones (type L)	[400–680]	564	Red
Cones (type M)	[400–650]	534	Green
Cones (type S)	[370–530]	420	Blue
Rods	[400–600]	498	Shade of gray

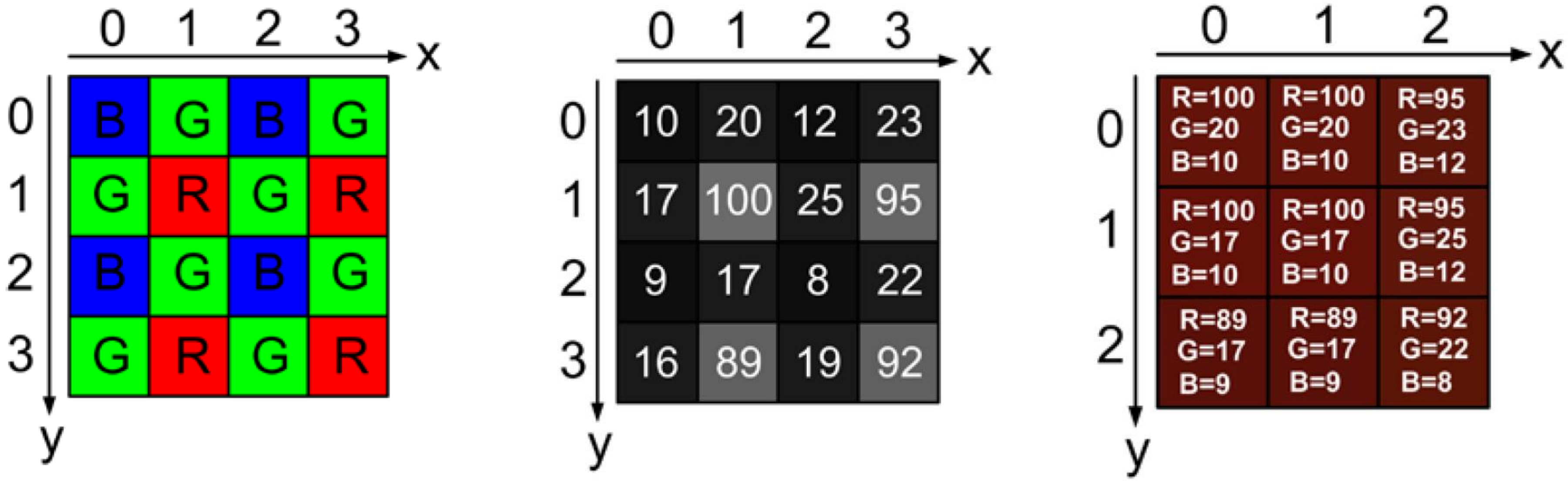
On average, a human perceives red as being 2.6 times as bright as blue and green as being 5.6 times as bright as blue. Hence the eye is more sensitive to green and least sensitive to blue.

RGB color space



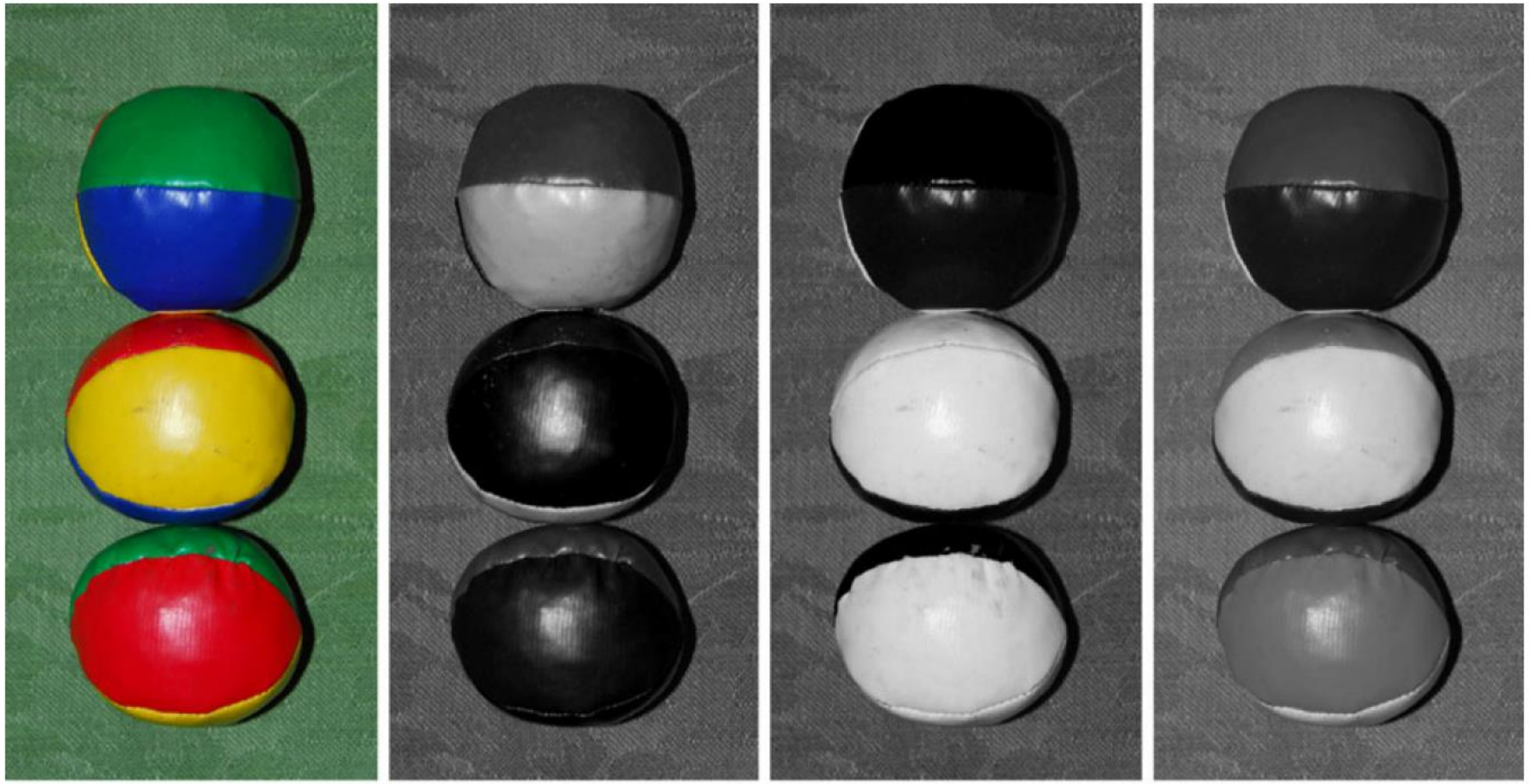
Corner	Color
(0, 0, 0)	Black
(255, 0, 0)	Red
(0, 255, 0)	Green
(0, 0, 255)	Blue
(255, 255, 0)	Yellow
(255, 0, 255)	Magenta
(0, 255, 255)	Cyan
(255, 255, 255)	White

Color image sensor (Bayer pattern)



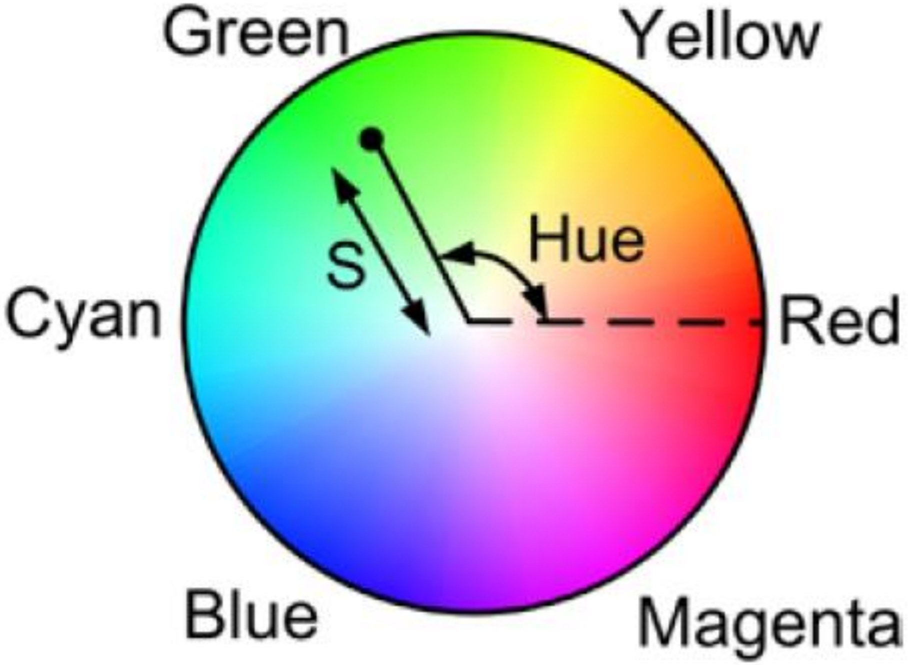
Converting from RGB to grayscale

- $I = W_R \cdot R + W_G \cdot G + W_B \cdot B; \quad W_R + W_G + W_B = 1$



Color input | $W_R = 0, W_G = 0, W_B = 1$ | $W_R = 1, W_G = 0, W_B = 0$ | $W_R = 0.5, W_G = 0.5, W_B = 0$

Alternative color representations (HSI)



RGB to HSI

$$H = \begin{cases} \cos^{-1}\left(1/2 \cdot \frac{(R-G)+(R-B)}{\sqrt{(R-G)(R-G)+(R-B)(G-B)}}\right), & \text{if } G \geq B; \\ 360^\circ - \cos^{-1}\left(1/2 \cdot \frac{(R-G)+(R-B)}{\sqrt{(R-G)(R-G)+(R-B)(G-B)}}\right), & \text{Otherwise} \end{cases}$$

$$H \in [0, 360[$$

$$S = 1 - 3 \cdot \frac{\min\{R, G, B\}}{R + G + B} \quad S \in [0, 1]$$

$$I = \frac{R + G + B}{3} \quad I \in [0, 255]$$