

Image Acquisition

F. Fotopoulos, Prof. Ioannis Pitas
Aristotle University of Thessaloniki
fototheo@csd.auth.gr
Version 5.2

Image Acquisition



- **Light Reflection**
- Camera Structure
- Camera Lens
- Sensor Technologies
- Image Digitization
- Image Corrections
- Image File Formats
- Scanners
- Image noise

Light Reflection



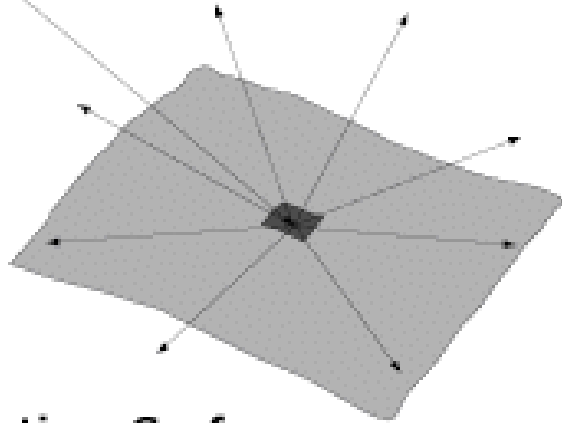
- Cameras capture objects that reflect or emit light.
- In general, reflection can be decomposed in two components:
 - ***Diffuse reflection***: Distributes light energy equally along any spatial direction.
 - ***Specular reflection***: Light energy is strongest along the direction of the incident light.
- Surfaces performing diffuse reflection, aka ***Lambertian surfaces***, are described as dull or matte; while specular reflection can be observed on glass surfaces and mirrors.

Light Reflection

Illumination



Reflected light

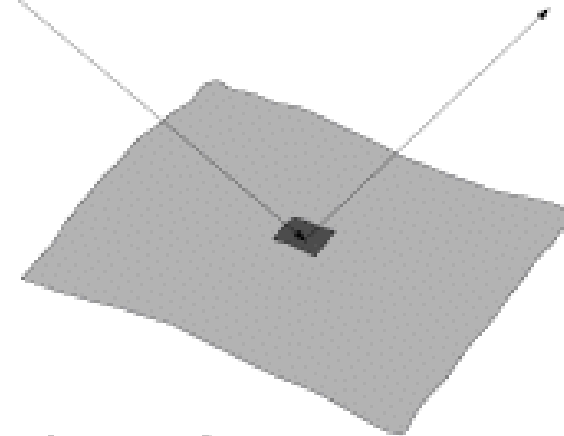


Lambertian Surface

Illumination



Reflected light



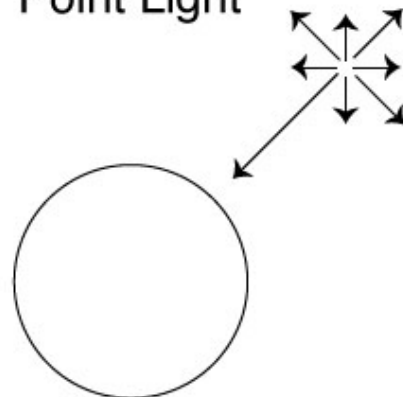
Specular Surface

Lambertian and specular surfaces [WOO2007].

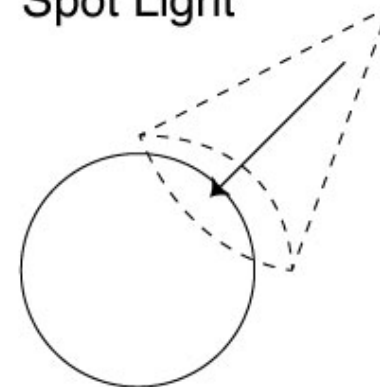
Light Reflection

- **Ambient illumination** sources emit the same light energy in all directions and thus, the position of the source is not that important for the reflecting surface (e.g., cloudy sky).
- **Point illumination** sources emit light energy isotropically (e.g., sun) or anisotropically (e.g., light bulb) along various directions making the position of the light source important.

Point Light



Spot Light



Light Reflection



- The reflected irradiance when object surface produces diffuse reflectance is given
- for **ambient illumination** by:

$$f_r(X, Y, Z, t, \lambda) = r(X, Y, Z, t, \lambda)E_a(t, \lambda)$$

- for **point light source** by:

$$f_r(X, Y, Z, t, \lambda) = r(X, Y, Z, t, \lambda)E_p(t, \lambda)\cos\theta.$$

- Distant point source and ambient illumination:

$$E(t, \lambda) = E_a(t, \lambda) + E_p(t, \lambda)\cos\theta.$$

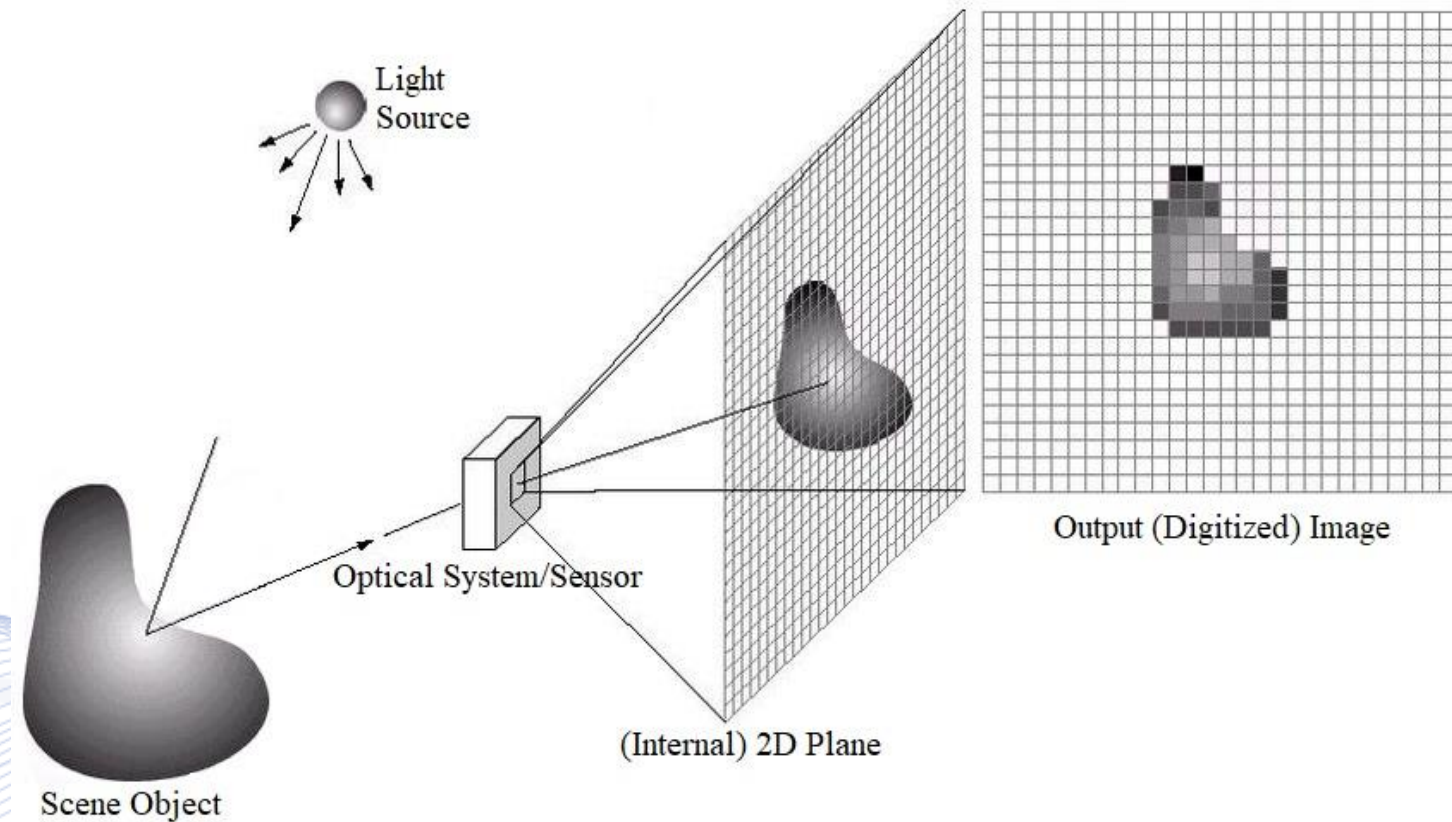
Image Acquisition



- We define an image as a **2D function** $f(x, y)$, where (x, y) corresponds to a co-ordinate in two-dimensional space and $f(x, y)$ represents the light intensity at this co-ordinate.
- In a discrete image, the (i, j) point is called a **pixel** (or picture element).
- Our main goal is to **project** a 3D scene into a 2D plane in which every point can be represented as a pixel.
- To achieve that we need an optical system that attracts the energy reflected from the object, as well as a **sensor** which computes the energy absorbed.

Image Acquisition

- The basic idea is that incoming illumination energy is converted into voltage and which can be subsequently sampled to obtain a digital image.
- CCD/CMOS sensor output varies according to the input electric energy.



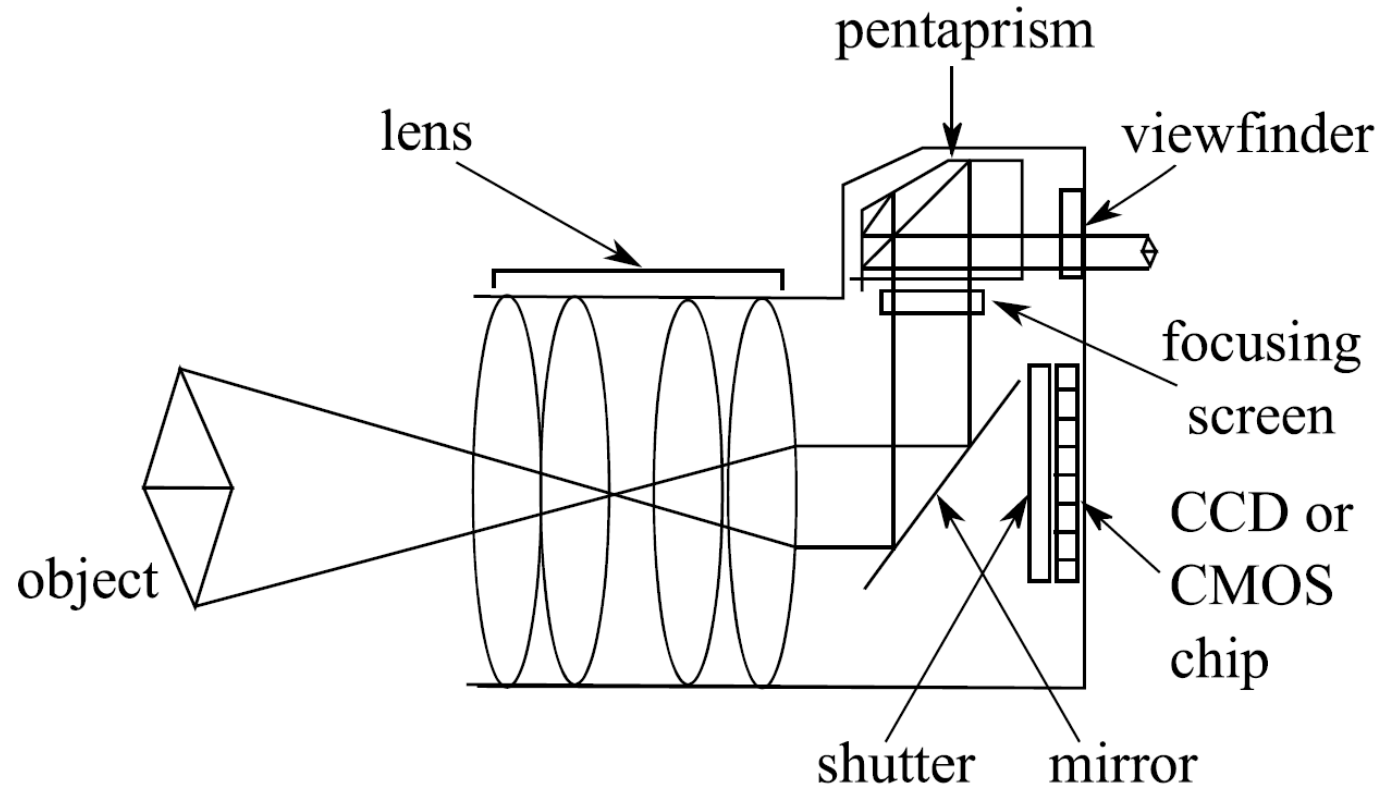
Basic Image Acquisition Model [MIS2017].

Image Acquisition



- Light Reflection
- **Camera Structure**
- Camera Lens
- Sensor Technologies
- Image Digitization
- Image Corrections
- Image File Formats
- Scanners
- Image noise

Camera Structure



Camera Exposure



- When it comes to exposure, the 3 basic camera terms are: aperture, shutter and ISO.
- **Aperture** refers to how “open” or “closed” the iris of a lens is, which affects the amount of light we let in.
- We measure aperture in **f-stops** and for larger f-stop numbers we get more closed iris which means that less light passes through.
- Also, by changing the f-stop we can control the depth of field.
- **Shutter** defines the time period that we are going to allow light rays to pass through and hit the camera’s sensor.

Camera Exposure



- High **shutter speed** is used when we want to capture fast moving scenes or get pictures without motion blur.
- **ISO** refers to sensor sensitivity to light.

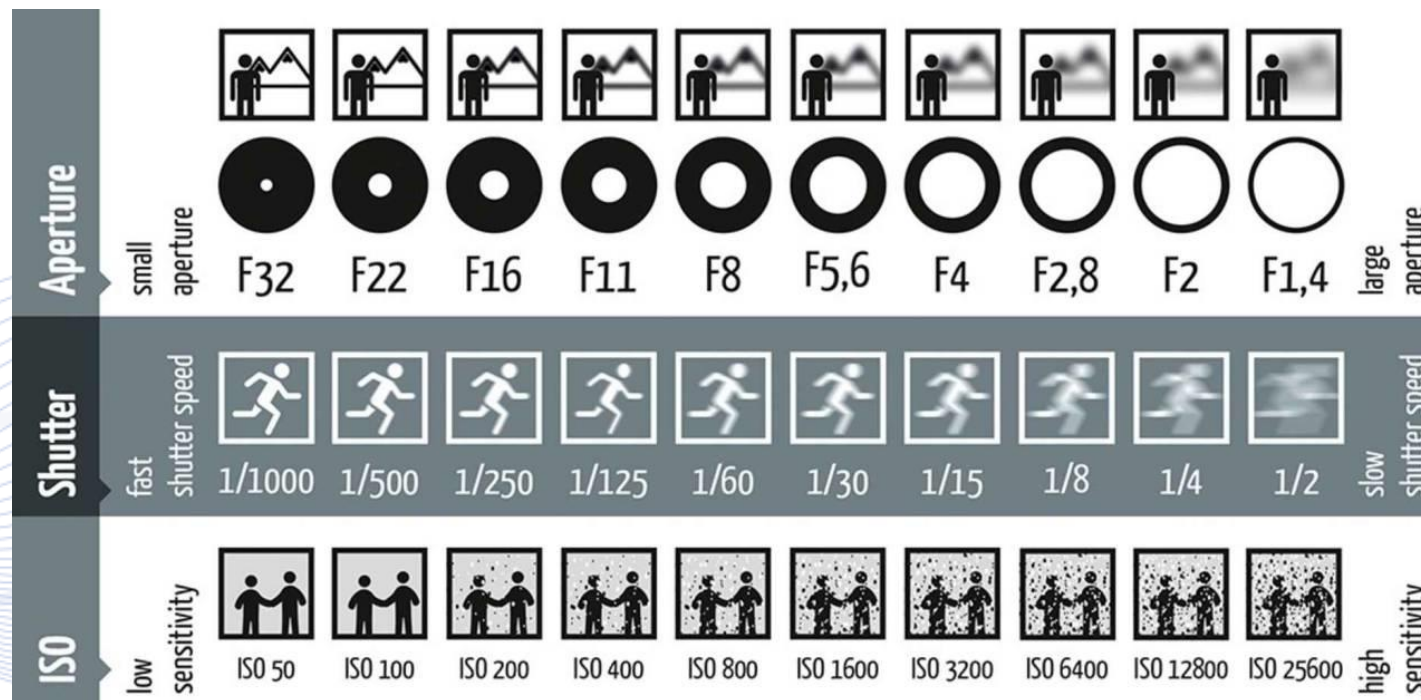


Image Acquisition



- Light Reflection
- Camera Structure
- **Camera Lens**
- Sensor Technologies
- Image Digitization
- Image Corrections
- Image File Formats
- Scanners
- Image noise

Camera Lens



- **Camera lens** are responsible for directing the light rays in order to recreate the scene as accurately as possible on the digital sensor.
- **Focal length** of a lens characterizes its angle of view and thus how much magnified will an object look for a given position.
- **Wide angle lenses** have short focal lengths, while **telephoto lenses** have longer focal lengths.
- **Zoom** lens gives you the ability to vary the focal length within a pre-defined range.
- A **prime or fixed** focal length lens cannot zoom.

Camera Lens



28mm wide-angle lens



300mm telephoto lens

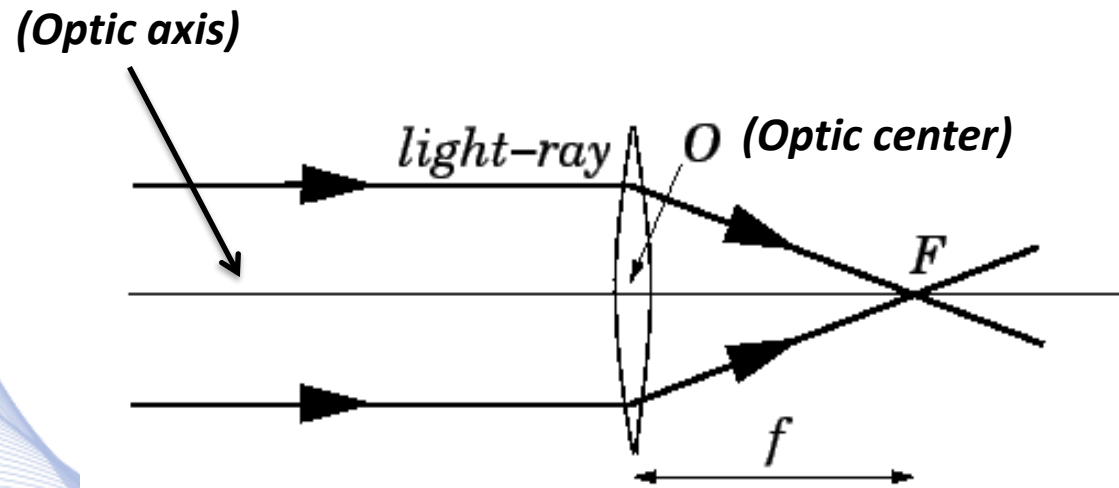
Wide-angle and telephoto lenses [PHO2013].

Optical Lens

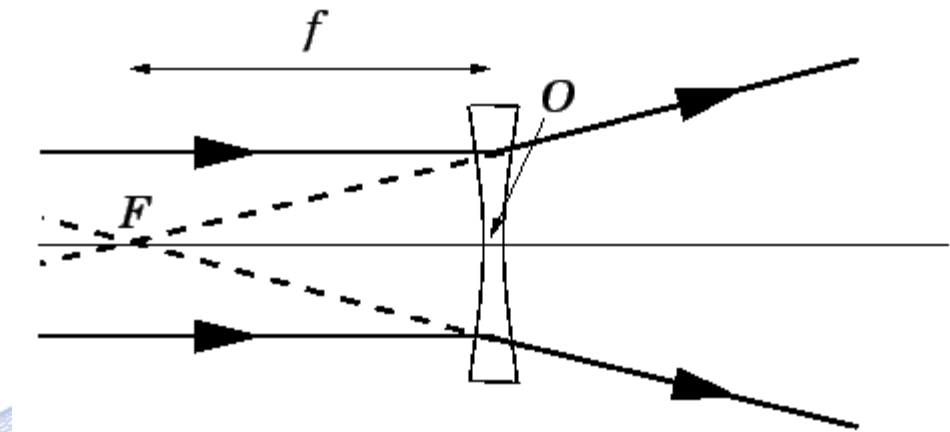


- A lens is a transparent medium (usually glass) bounded by two curved surfaces.
- The two major kinds of lenses are converging and diverging.
- A **converging lens** brings all light-rays parallel to optical axis together at the so-called focal point **F**.
- A **diverging lens** spreads out all light-rays parallel to optical axis in such a way that they appear to diverge from a virtual focal point **F** in front of the lens.

Optical Lens



a) Converging Lens



b) Diverging Lens

Optical Lens



- ***Lensmaker's equation*** gives us the focal length (f):

$$\frac{1}{f} = (n - 1) \left(\frac{1}{R_1} - \frac{1}{R_2} \right).$$

- n : index of refraction
- R_1, R_2 : radii of curvature of the two surfaces

- Thin lens equation:

$$\frac{1}{d} + \frac{1}{d'} = \frac{1}{f}$$

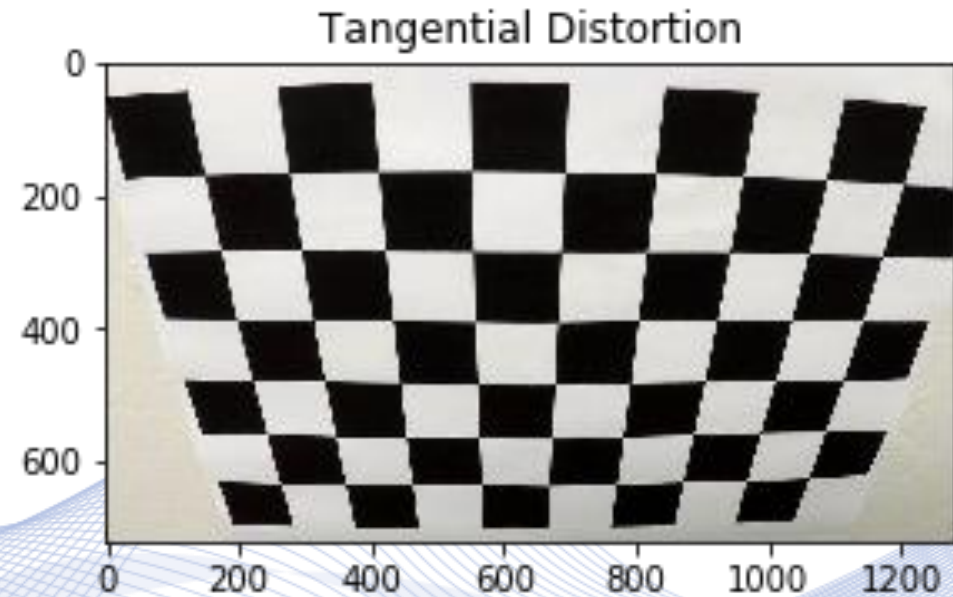
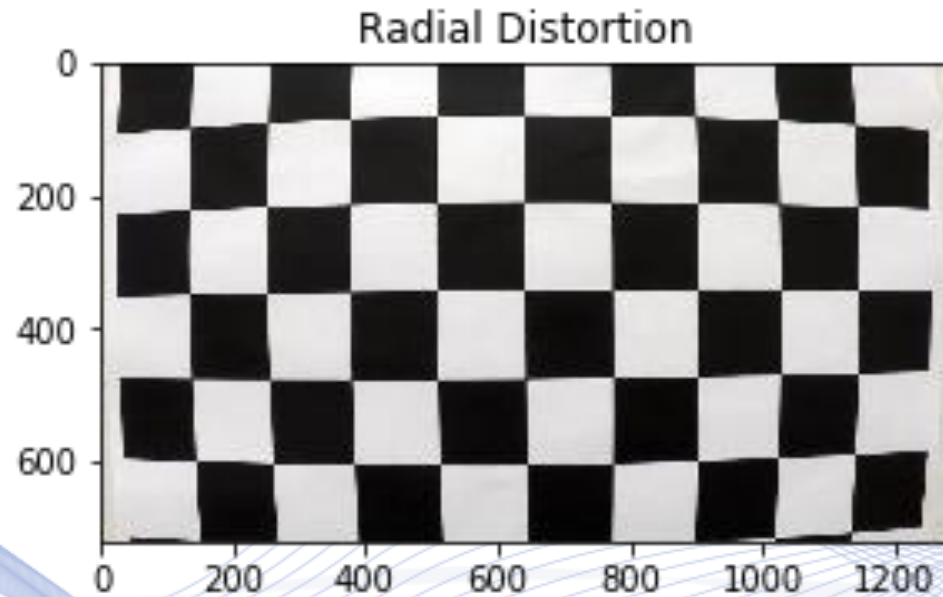
- d : object distance
- d' : image distance

Lens Distortion



- A common problem that occurs when trying to produce a digital image via a camera is lens distortion.
- This causes physical straight lines to appear as curves in the output image.
- The two main categories of lens distortions are:
 - ***Radial distortion.***
 - ***Tangential distortion.***

Lens Distortion



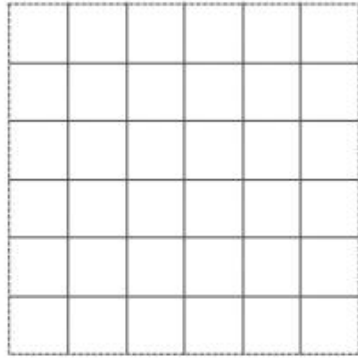
Lens Distortion



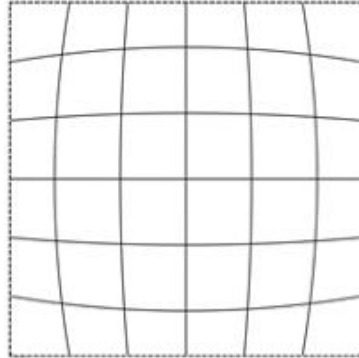
Radial distortion

- It is caused by the geometric shape of the lens and affects the imaging of straight lines.
- Most often, radial distortion affects to a much greater degree, the geometry of the image than tangential distortion.
- The types of radial distortions are:
 - *Barrel Distortion.*
 - *Pincushion distortion.*
 - *Mustache Distortion.*

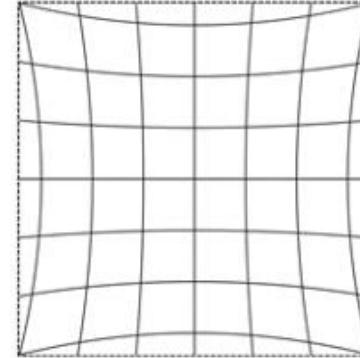
Lens Distortion



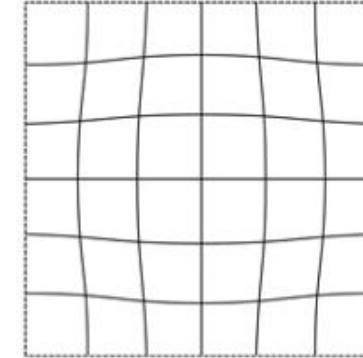
no distortion



barrel distortion



pincushion distortion



mustache distortion

- We can observe that radial distortions are symmetric about the image center.
- Barrel distortion is found in wide-angle/panoramic images, while pincushion distortion is associated with telephoto lenses to eliminate globe effects.

Lens Distortion

- To correct radial distortion, we use:

$$x_{ud} = x_d(1 + k_1r^2 + k_2r^4 + k_3r^6),$$

$$y_{ud} = y_d(1 + k_1r^2 + k_2r^4 + k_3r^6),$$

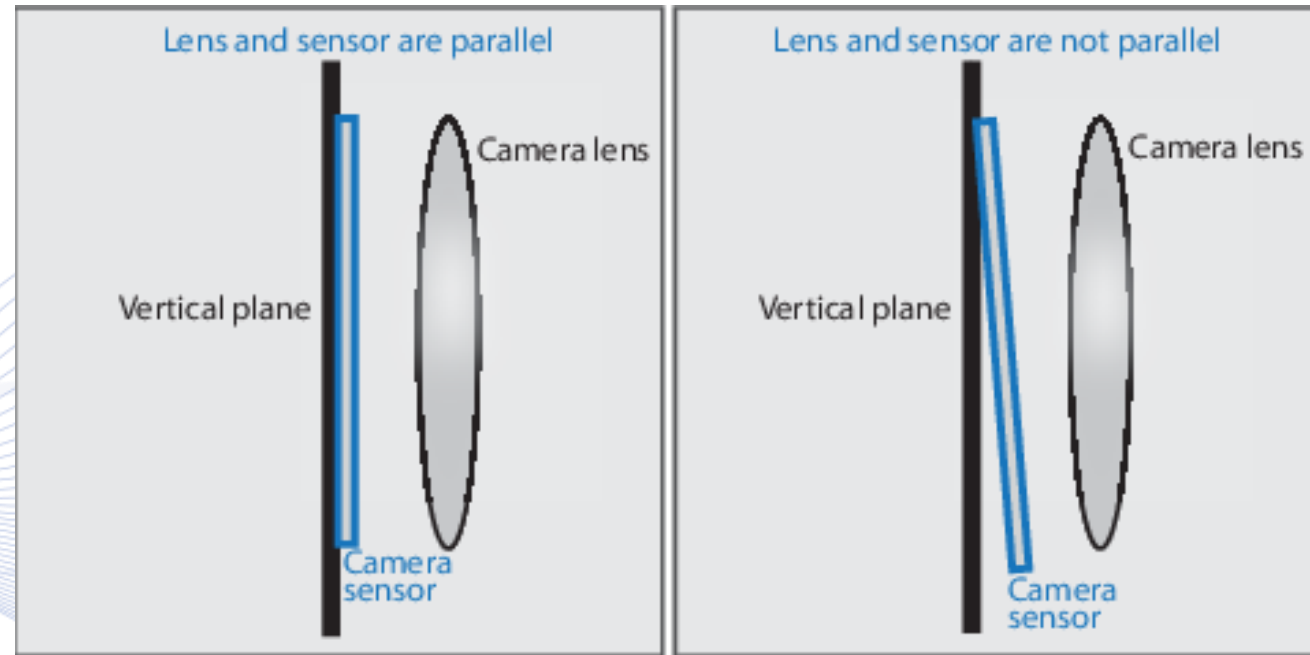
$$r^2 = x_d^2 + y_d^2.$$

- where $x_d = \frac{x^c}{z^c} = \frac{u-u_0}{f_x}$ and $y_d = \frac{y^c}{z^c} = \frac{v-v_0}{f_y}$ are obtained from the distorted image.
- k_1 , k_2 and k_3 are three parameters used for radial distortion correction.

Lens Distortion

Tangential distortion:

- It occurs when the lens installed is not perfectly parallel to the image plane.



(3.1) Source: Mathworks

Lens Distortion



- The correction formula for tangential distortion is:

$$\begin{aligned}x_{ud} &= x_d + [2p_1x_dy_d + p_2(r^2 + 2x_d^2)], \\y_{ud} &= y_d + [p_1(r^2 + 2y_d^2) + 2p_2x_dy_d].\end{aligned}$$

- where p_1 and p_2 are two intrinsic parameters that can be estimated using the appropriate collection of images.

Image Acquisition

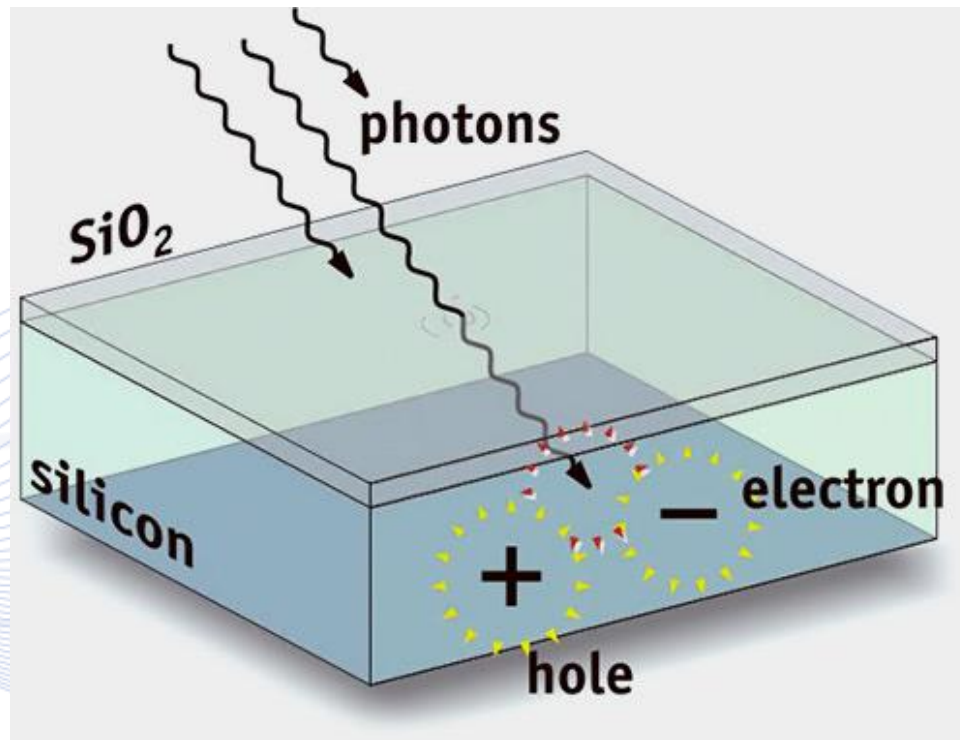


- Light Reflection
- Camera Structure
- Camera Lens
- **Sensor Technologies**
- Image Digitization
- Image Corrections
- Image File Formats
- Scanners
- Image noise

Sensor Technologies



- **CCD** and **CMOS** are the two most common sensor technologies that digital cameras use are.
- Both convert light energy into electrons.



Light sensing [LIT2001] .

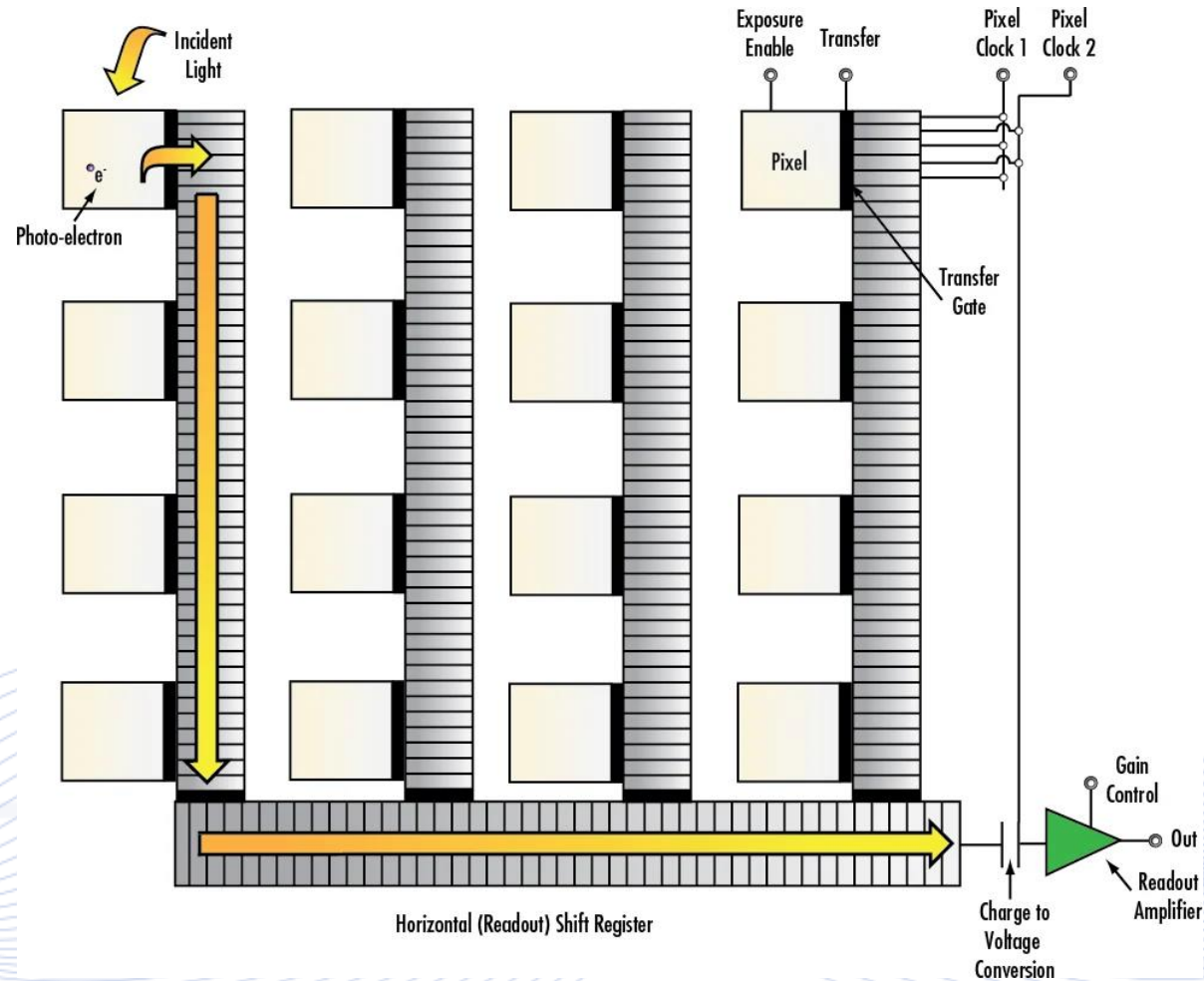
Sensor Technologies



Charge-Coupled Device (CCD)

- A silicon chip that holds an array of photosensitive areas.
- Charged packets are transferred from these areas into a shift register.
- Ultimately, electron energy passes into an analog-to-digital (A/D) converter which outputs the pixel value.
- Efficiency in speed relies on the advance of charged packets across the chip. Thus, it is limited.
- Yet, this phenomenon contributes to the high sensitivity and pixel-to-pixel consistency of the CCD (uniformity).

Sensor Technologies



CCD architecture [EDM2021].

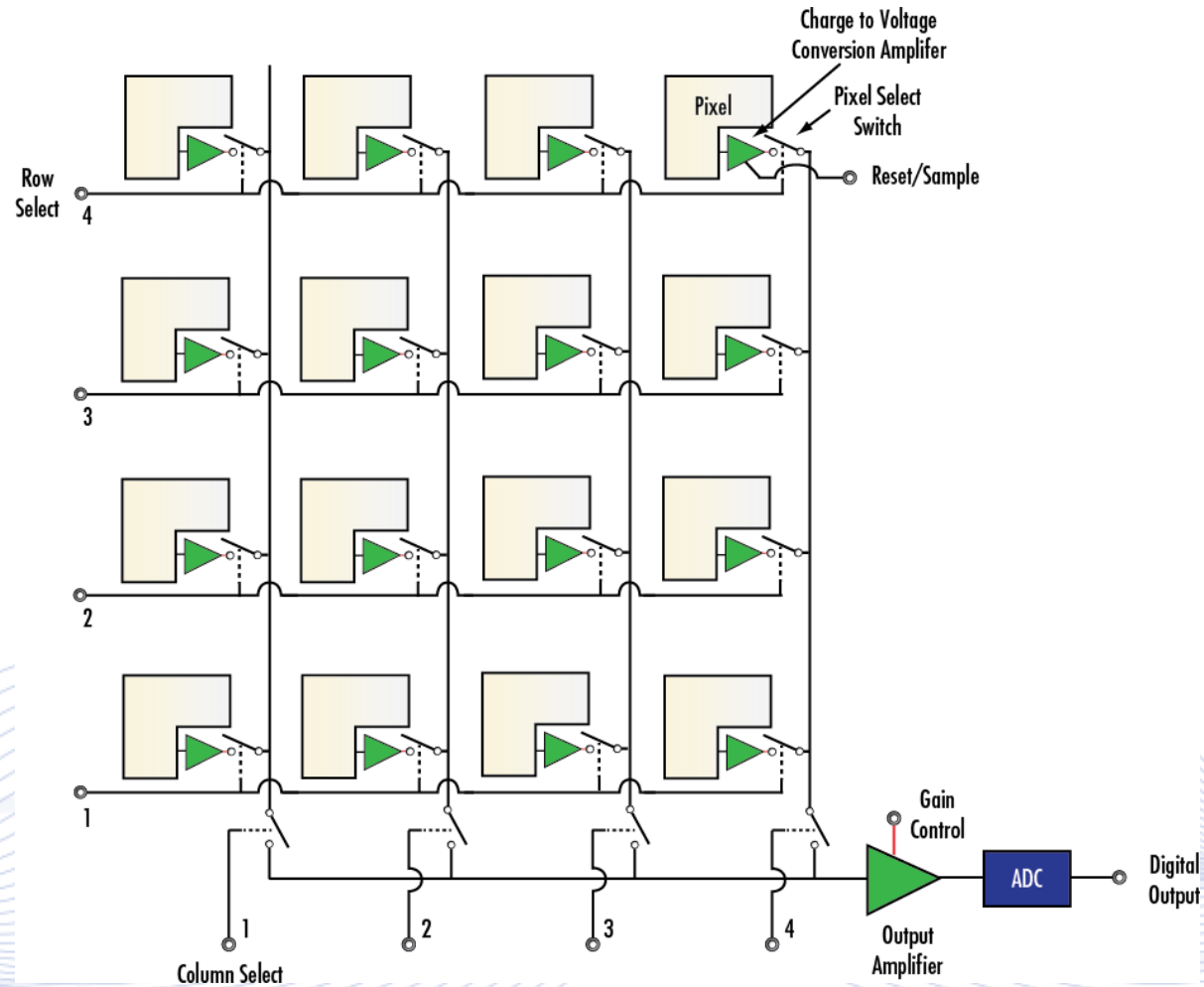
Sensor Technologies



Complementary Metal Oxide Semiconductor (CMOS)

- The charge from the photosensitive pixel is converted to a voltage at the pixel site and gets amplified
- The signal is multiplexed by row and column with the rest of the chip.
- Each pixel has its own charge-to-voltage conversion, and often also contains amplifiers, noise-correction, and digitization circuits
- This parallel process allows for high bandwidth/speed but decreases its uniformity.

Sensor Technologies



CMOS architecture [EDM2021].

CCD VS CMOS



- **Responsivity** is the amount of signal the sensor delivers per unit of input energy. CMOS are slightly superior to CCD.
- **Dynamic range** is the ratio of a pixel saturation level to its signal threshold. CCDs gives better results.
- **Uniformity** is the consistency for different pixels under the same illumination circumstances.
 - CMOS sensors are traditionally worse in that area, but latest CMOS technologies come close to the CCD.
- **Speed.** CMOS has bigger speed than CCD.

Image Acquisition



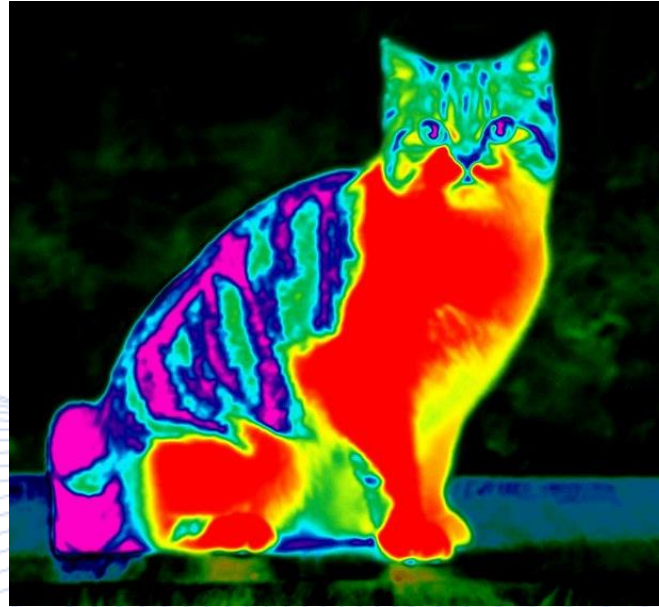
- Light Reflection
- Camera Structure
- Camera Lens
- Sensor Technologies
- **Image Digitization**
- Image Corrections
- Image File Formats
- Scanners
- Image noise

Image Types

- The image acquired depends on the type of camera (and film) that we use.



a) X-ray image



b) Infrared image



c) Visible light image

Image Types

The three basic types of images are:

- **Binary images:** Each pixel can take a value of 0 or 1: $g(x, y) \in \{0,1\}$.
- **Gray-scale images:** Each pixel can take a value from 0 to 255:

$$f(x, y) \in \{0, \dots, 255\}$$

- **Color images:** Each pixel consists of three channels (RGB). Each of them is assigned with a value from 0 to 255: $[f_R(x, y), f_G(x, y), f_B(x, y)]^T \in \{0, \dots, 255\}^3$.



Image Digitization



- As we previously seen the sensor produces an internal 2D plane which is basically a projection of the 3D scene that has been captured.
- In order to represent that 2D plane digitally in a computer we need to go through the process of digitization.
- Specifically, this consists of two basic subprocess:
 1. ***Spatial sampling*** of the 2D plane.
 2. ***Quantization*** of the light intensity values that has been taken via the sensor.

Quantization

- Each point in the 2D plane has a light intensity value that ranges from 0 to ∞ , so it must be quantized.
- Usually we produce 8-bit images, so for example in a gray-scale image we can use up to $256 = 2^8$ gray levels.



a) 256 gray levels



b) 64 gray levels



c) 16 gray levels



d) 2 gray levels

- Correspondingly, in RGB values should be quantized for each channel.

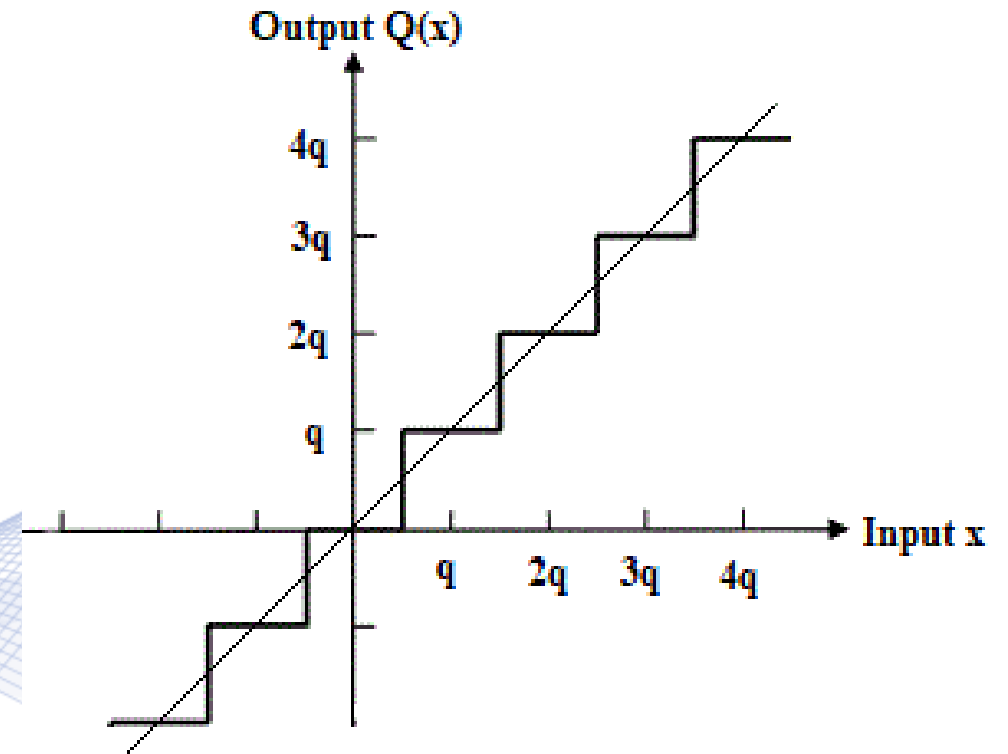
Quantization

- For uniform distribution, we define q as the quantization step:

$$q = \frac{1}{2^b}.$$

- for 8-bit images we have $b = 8$.
- And to quantize the light intensity levels:

$$kq, \quad k = 0, 1, 2, \dots, 2^b - 1.$$



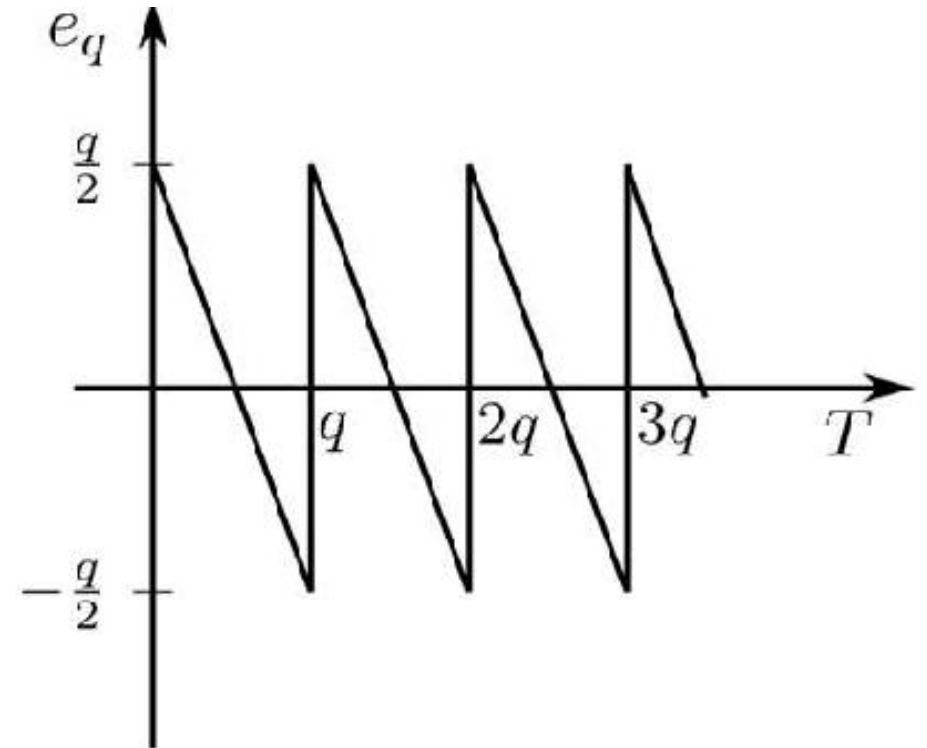
Quantization

- **Mean square Error (MSE)** of quantization:

$$e_q = E \left[(x - Q(x))^2 \right] = \int_{-\infty}^{\infty} (x - Q(x))^2 p(x) dx.$$

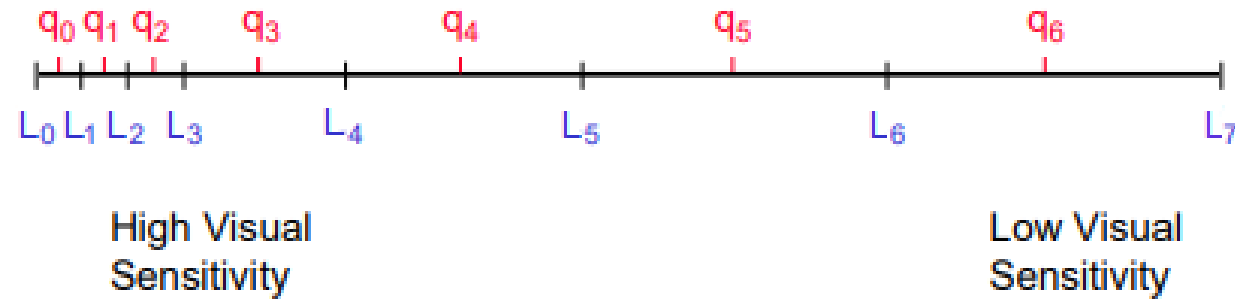
- $p(x)$ is distribution of input x .
- **Signal-to-Noise Ratio (SNR)** is given by:

$$SNR = 20 \log_{10}(2^b) \approx 6.02b \text{ dB}.$$



Non-uniform Quantization

- Using Weber's law, we can quantize regarding visual sensitivity:



- We can also quantize according to sensor output:

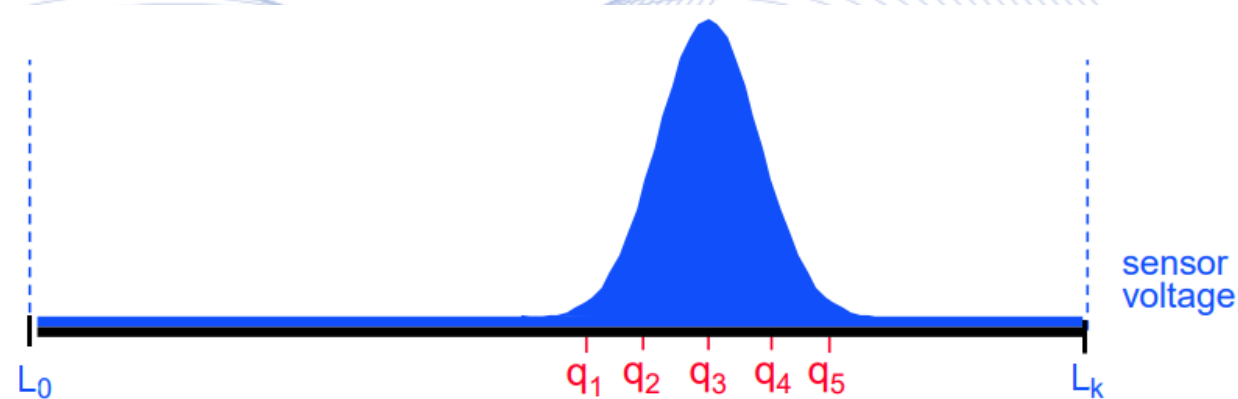


Image Acquisition



- Light Reflection
- Camera Structure
- Camera Lens
- Sensor Technologies
- Image Digitization
- **Image Corrections**
- Image File Formats
- Scanners
- Image noise

Gamma (γ) Correction



- Gamma defines the relation between the given pixel value and its actual luminance.
- Cameras do not capture luminance in a linear fashion so in order to correct that we use the gamma correction:

$$v_{out} = v_{in}^{\gamma}.$$

- Image displays also suffer from non-linear light intensity depiction.
- Most common Gamma used are 2.2, 2.4, 2.6.

Gamma (γ) Correction



Gamma Correction [ROS2021].

Contrast equalization



- **Just Noticeable Difference (JND)** is the minimum amount something must be changed for the difference to be apparent.
- Perceived contrast is a function of the intensity, specifically via **Weber's law** we get:

$$\frac{\Delta f}{f} = k.$$

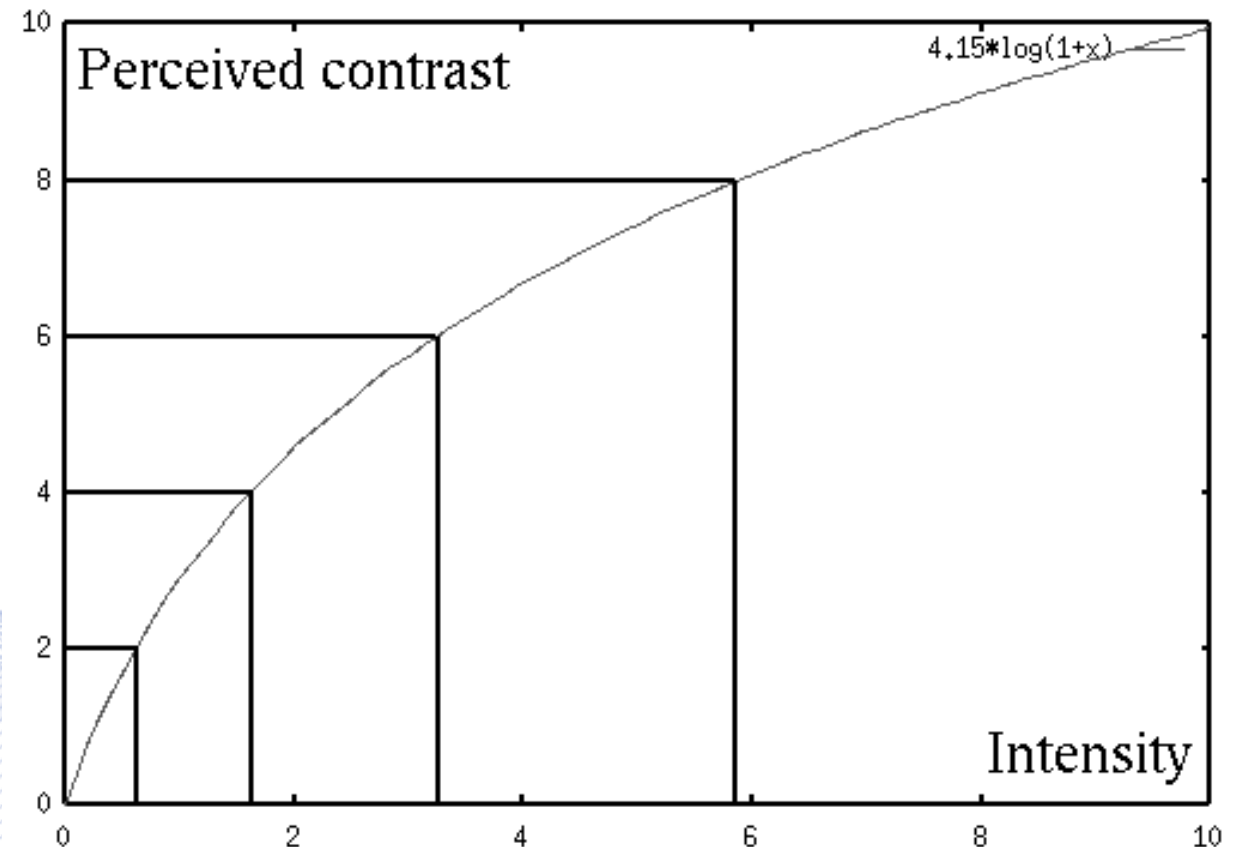
where f is the intensity and Δf is the intensity difference (contrast).

- This implies that sometimes human eye can not perceive the contrast between different intensities on an object and its background.

Contrast equalization

- For example, a high contrast $\Delta f = 200 - 100$ at a background intensity $f = 100$ is observed the same way as lower contrast $\Delta f = 20 - 10$ with $f = 10$ (because they both equal 1).
- This basically means that we are less sensitive when the intensity level is high.
- **Perceived contrast equation:**

$$y = a \log(1 + x).$$



White balance

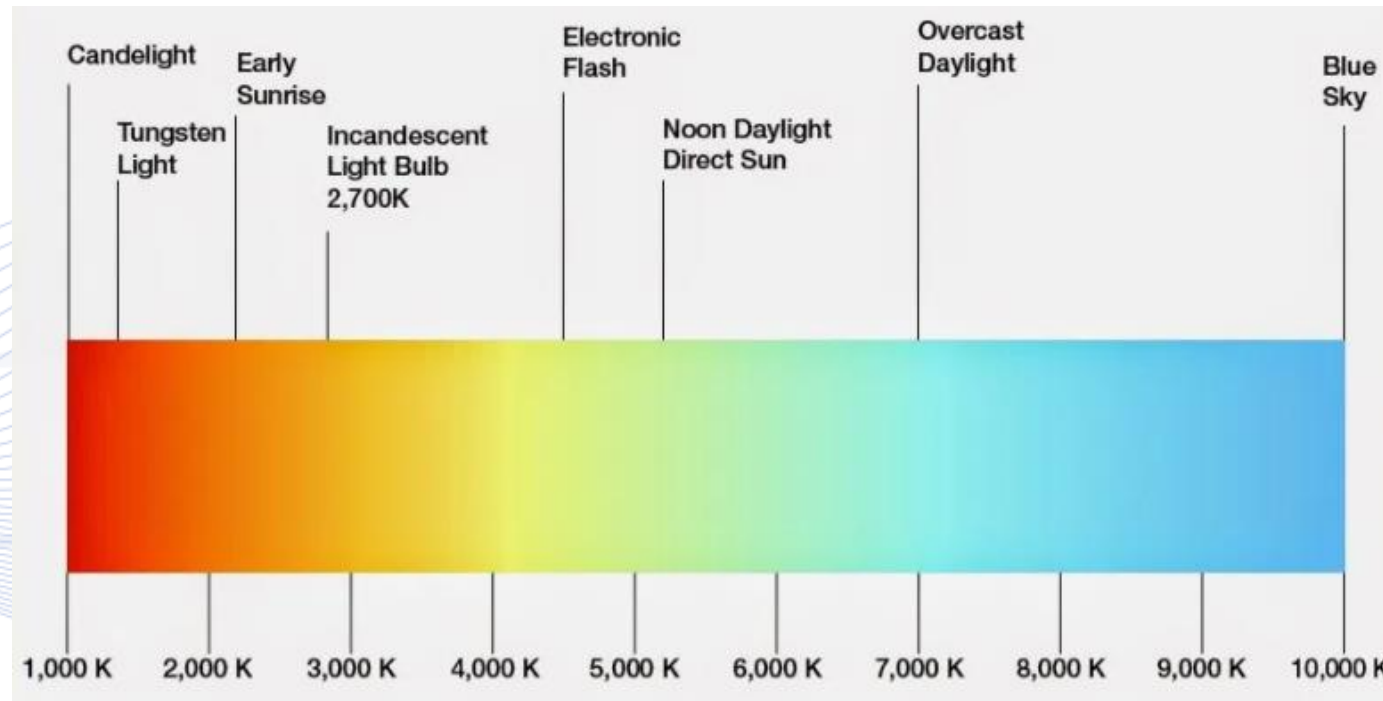
White balance is the process of removing unrealistic colors that are cast into the image, in order to achieve natural ones.

- To achieve that we must consider the color temperature of the light source.



White balance

- We measure **color temperature** in Kelvin (K) and is a physical feature of light.
- As the color temperature rises, we move from warm colors to cooler ones.



White balance



White balance in photography [WIL2021].

White balance

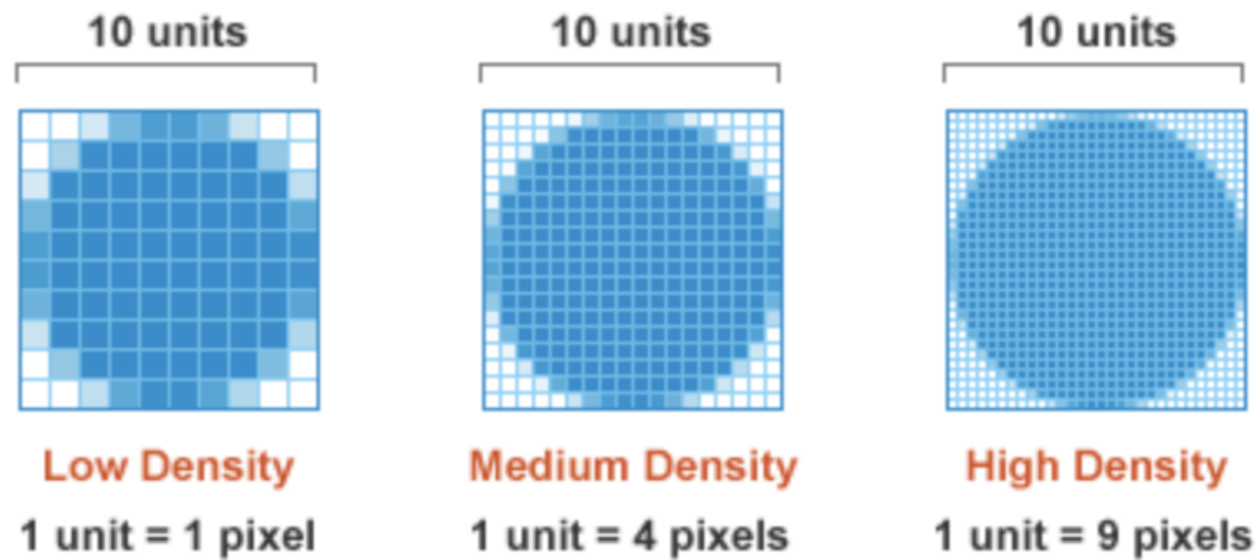
- To balance the colors in a picture we need to scale the relative luminances, by using as reference a white/neutral object in the image.

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 255/R'_w & 0 & 0 \\ 0 & 255/G'_w & 0 \\ 0 & 0 & 255/B'_w \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix}.$$

- where R , G and B are the color balanced RGB values of a pixel in the image; R' , G' and B' are the RGB components before color balancing and R'_w , G'_w , B'_w corresponds to the RGB values of a pixel that is believed to be a white surface.

Spatial Sampling

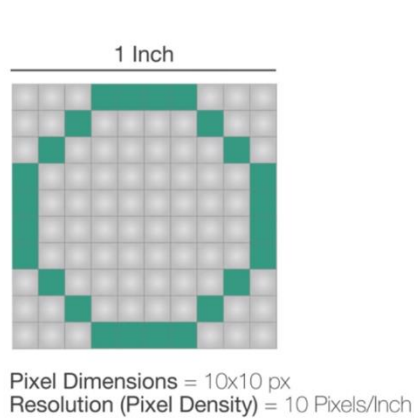
- The continuous 2D plane needs to be spatially sampled in order to be rasterized.
- The more pixels a rasterization unit has the best of a resolution we can achieve.



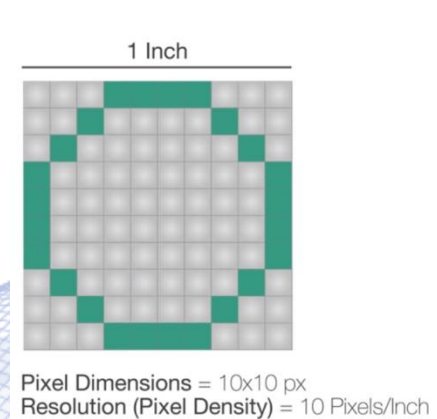
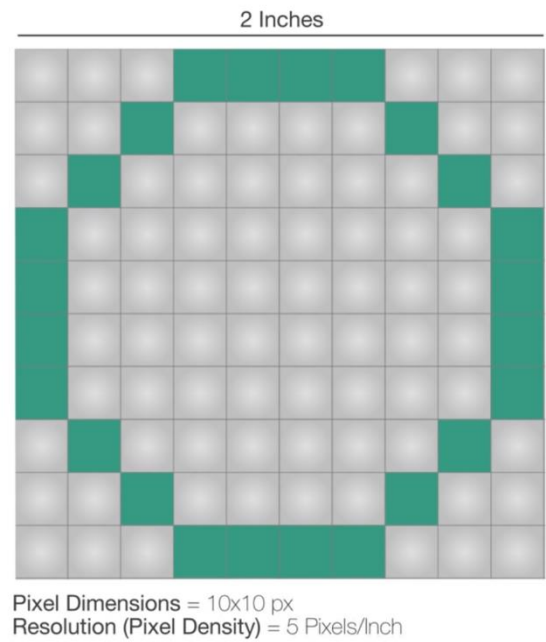
Different pixel densities [RAH2021].

Image Resolution

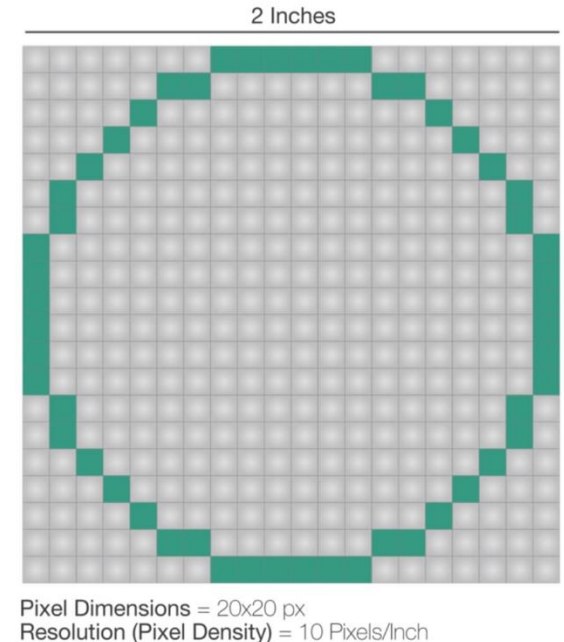
- Two frequently misunderstood terms are *image size* and *image resolution*.
- Image size refers to **pixel dimensions** of the image while image resolution refers to **pixel density (PPI)**. (Note: DPI≠PPI)



➤ *If we increase the physical size (inches) and keep the same dimensions (pixels) we have lower resolution.*



➤ *We need to increase the pixel dimensions when physical size gets bigger in order to preserve the resolution.*



Sampling Rate



- **Aliasing** occurs when sampling a continuous signal/image the sampling rate isn't high enough and results to faulty capture of the image.
- From **Nyquist rule**, the image sampling rate should be greater than double the maximal spatial image frequency in the image to avoid aliasing:

$$f_{sx} > 2f_{xmax}, \quad f_{sy} > 2f_{ymax}.$$

- Problematic sampling process can lead to picture artefacts, like Moire patterns.

Sampling Rate

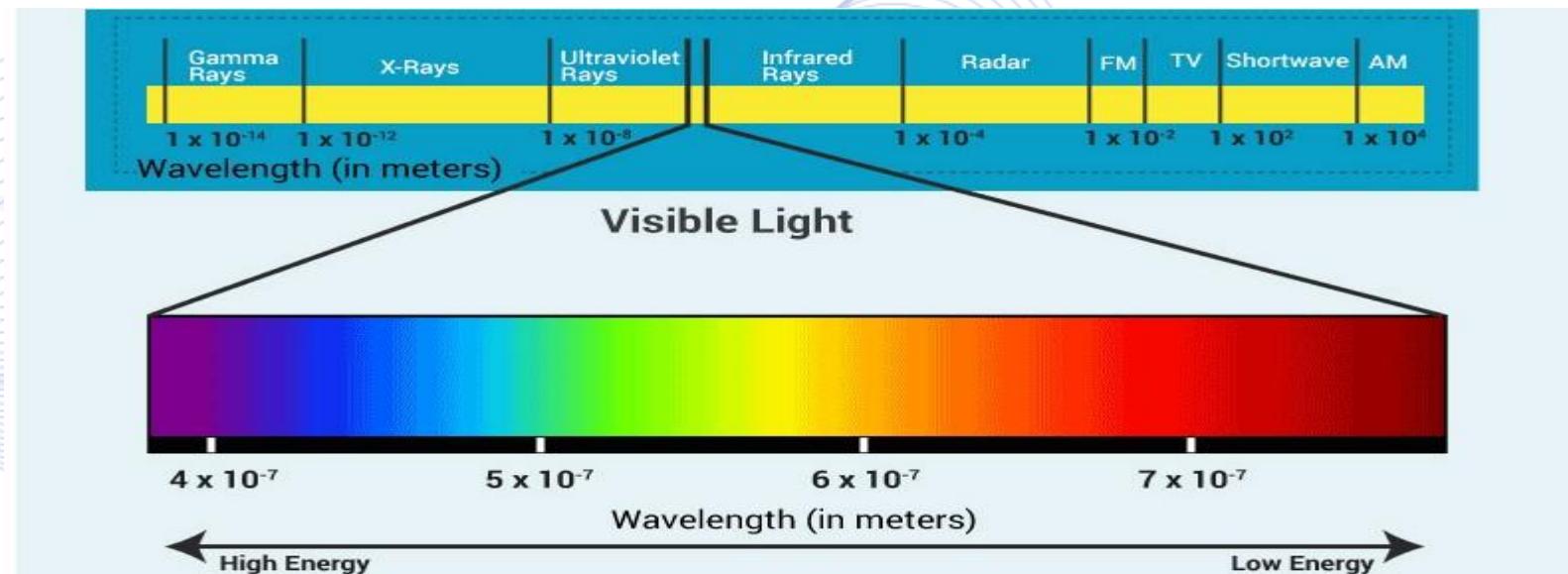
- Examples of Moire patterns:



Demosaicing

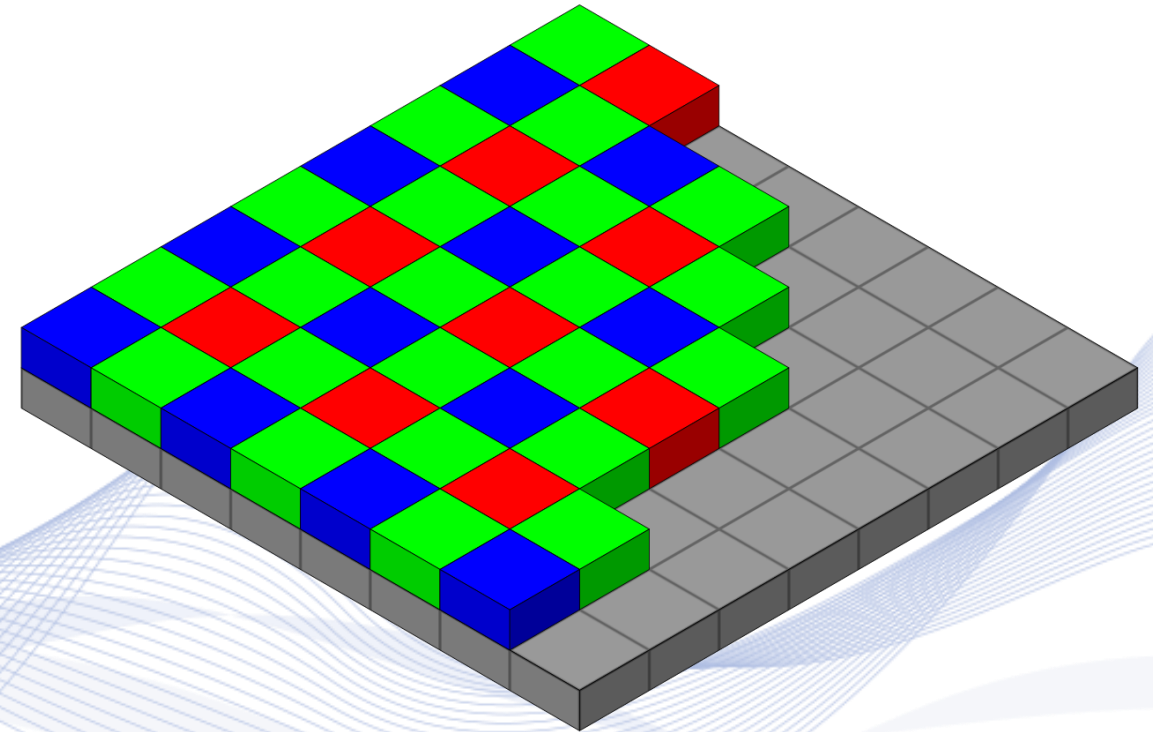


- **Color filter array** or **color filter mosaic** is an array consisting of tiny color filters that are placed in front of the camera sensor.
- They are essential for capturing RGB images because sensor by themselves can only produce gray-scale images.
- Each color filter allows different wavelength range of light to pass in order to obtain information about the color of light.



Demosaicing

- The most common type of Color filter array is the **Bayer filter**.
- The filter patterns consists of 50% green, 25% red and 25% blue.
- This happens because human eyes are more sensitive to green and distinguish luminance intensity better in green channel than in red or blue.



Bayer Filter [BAY2021].

Demosaicing



- To construct the full RGB image an algorithm is implemented which interpolates the missing color values for all the pixels.
- Simple algorithms that implement demosaicing interpolation are:
 - **Nearest neighbor interpolation:** simply copies the closest pixel of the same color channel (low quality method)
 - **Bilinear interpolation:** where for example a blue value is computed for a non-blue pixel as an average of the four nearest blue color pixels (or two nearest if it's linear interpolation)
- Better quality methods include bicubic interpolation, spline interpolation.

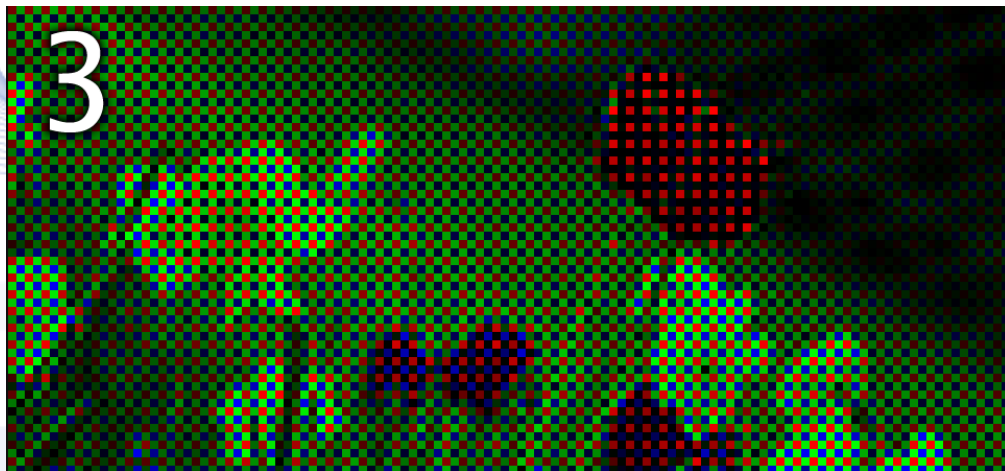
Demosaicing



1) Original scene



2) Gray-scale output of sensor array



3) Output array via Bayer filter



4) Reconstructed image after interpolation

Image Acquisition



- Light Reflection
- Camera Structure
- Camera Lens
- Sensor Technologies
- Image Digitization
- Image Corrections
- **Image File Formats**
- Scanners
- Image noise

Image File Formats



- The types of formats that an image data can be stored are:
 - **Uncompressed format:** Contains all the data without any compression applied. This results in very big image sizes. (TIFF, RAW, BMP)
 - **Compressed format:** Compression algorithms are applied, 2 subcategories:
 - **Lossless:** File size is reduced without losing any of the uncompressed image data. (PNG, GIF, FLIF)
 - **Lossy:** File size is reduced but data from the uncompressed image are thrown away which can lead to noise/artifacts appearing in the image. (JPEG)

Image File Formats

- Vector format:** Unlike the previous ones, vector format (SVG, CGM) does not apply rasterization, instead it contains a geometric description of the image. Therefore, it can be displayed smoothly at any desired size.

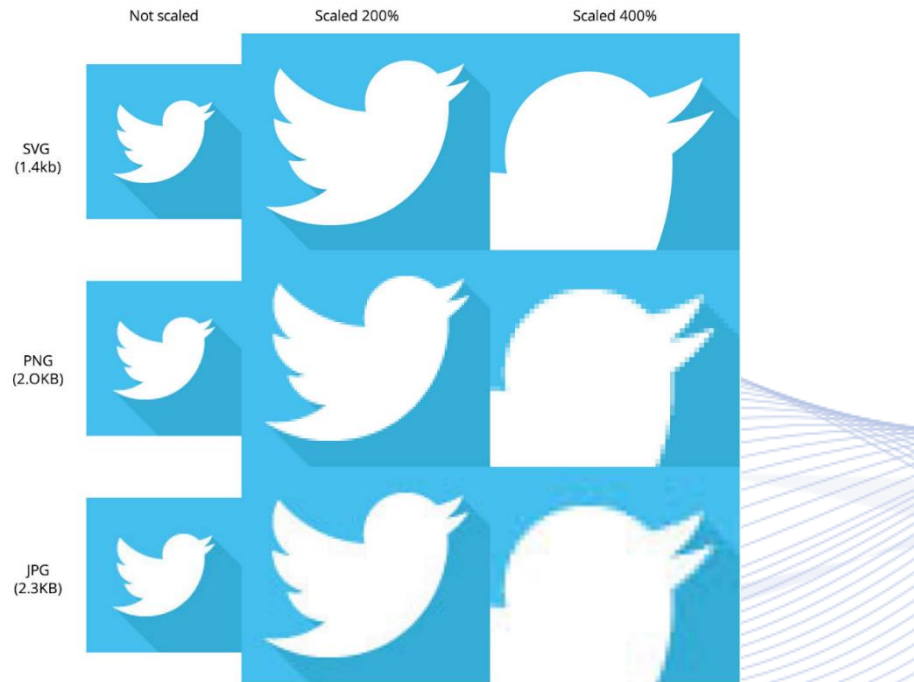


Image Acquisition



- Light Reflection
- Camera Structure
- Camera Lens
- Sensor Technologies
- Image Digitization
- Image Corrections
- Image File Formats
- **Scanners**
- Image noise

Scanners



- Most common types of scanners are:
 - **Flatbed scanners:** Most versatile and commonly used scanners.
 - **Sheet-fed scanners:** Similar to flatbed scanners except that the document is moving and the scan head is fixed.
 - **Handheld scanners:** Same basic technology as flatbed scanner but rely on the user to move them. Typically, mediocre quality but is a quick-scan method.
 - **Drum scanners:** Mostly used by publishing industry to obtain incredibly detailed images.

Scanners



a) Flatbed scanner



b) Sheet-fed scanner

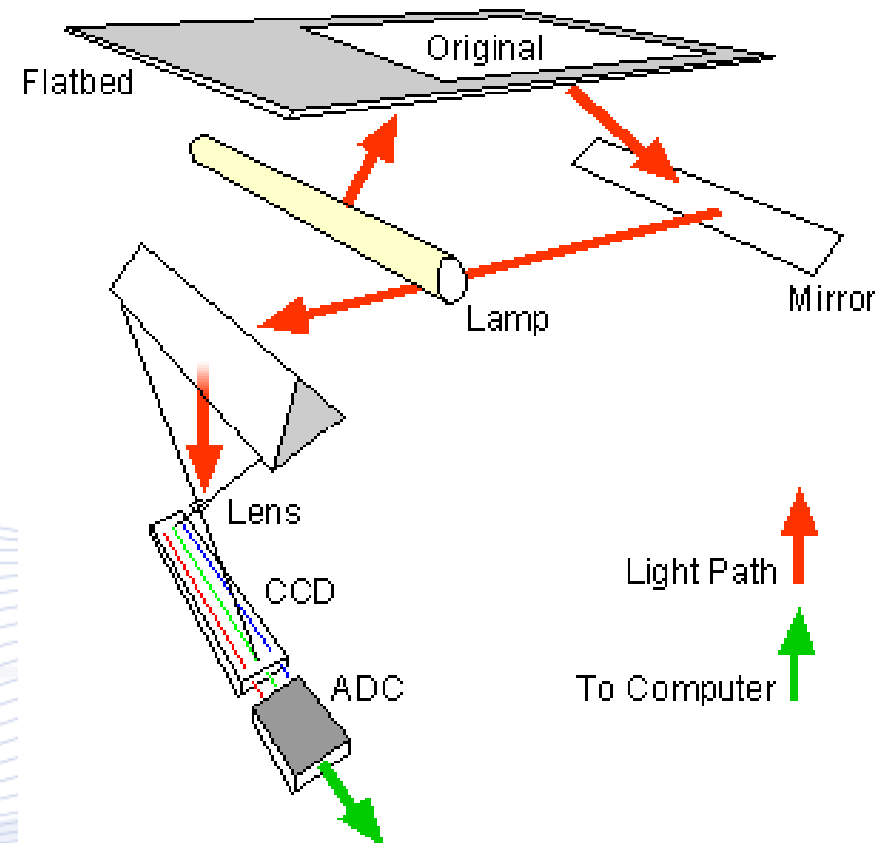


c) Handheld scanner



d) Drum Scanner

Scanners



Anatomy of the flatbed scanner [COP2021].

Scanners



Scanning process:

- Document is placed on the glass plate and the **scan head** (mirrors, lens, CCD) is placed in the starting position.
- Lamp illuminate the sheet of paper and reflects back to the **mirror** while the scan head is moving slowly across the document.
- The image of the document is reflected from the first **angled mirror** to another mirror and then finally onto a **lens**.
- The lens focuses the image through a filter on the **CCD**.
- Finally, A/D converter is used to get the digital output of CCD.

Image Acquisition



- Light Reflection
- Camera Structure
- Camera Lens
- Sensor Technologies
- Image Digitization
- Image Corrections
- Image File Formats
- Scanners
- **Image noise**

Image noise

- Image noise can be mainly of two kinds:
 - **Photoelectronic noise:** Generated by the image sensor and the circuits inside the digital camera.
 - **Film-Grain noise:** Optical texture that appears mainly in processed photographic film. It can be added on deliberately for stylistic purposes or to make a digital image look less “flat”.
- Most common noise types observed in digital images are Gaussian noise, Poisson noise, Speckle noise and Salt-and-Pepper noise.

Image noise



(Digital Noise)



(Film Grain)

Image noise

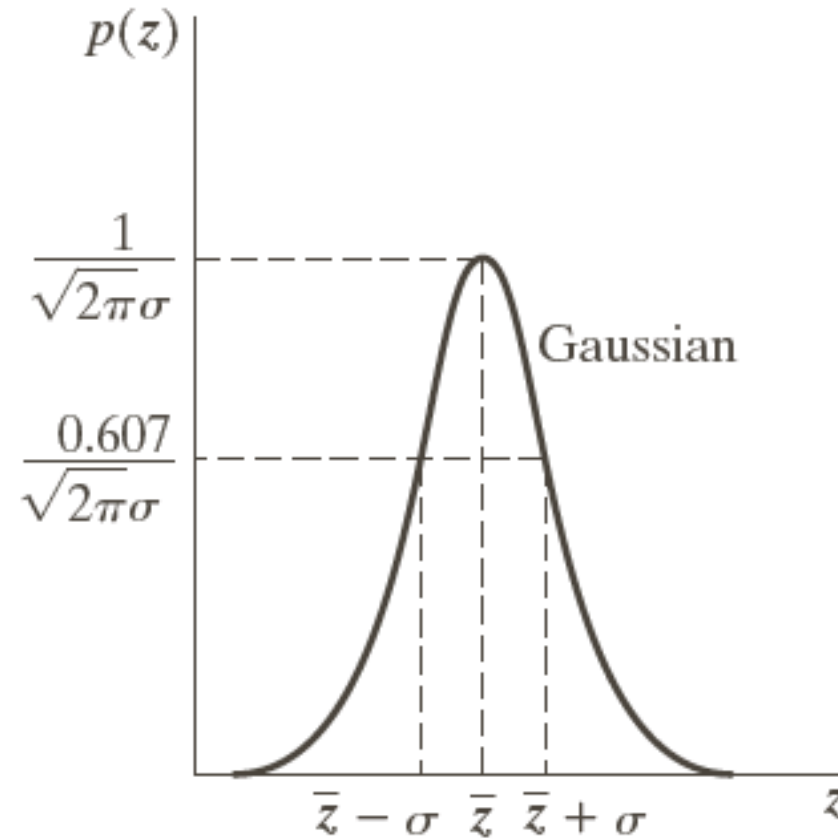
Gaussian noise:

- Arises in amplifiers or detectors and is caused by natural sources (e.g. thermal vibration of atoms).
- Gaussian noise mainly distorts gray values in the image.
- PDF of gaussian noise is given by:

$$p(f) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(f-\mu)^2}{2\sigma^2}}$$

- f represents the gray level, μ the mean and σ the standard deviation

Image noise



Gaussian noise PDF.

Image noise

Salt-and-Pepper noise:

- Usually caused by sharp and irregular disturbances in the image signal.
- Also called impulse noise because it consists of black and/or white impulses which alter random pixels in the image.
- Model:

$$g(i, j) = \begin{cases} z(i, j), & \text{with probability } p \\ f(i, j), & \text{with probability } 1 - p. \end{cases}$$

Image noise



Speckle noise:

- Multiplicative type of noise that inherently exists in systems such as active radar and ultrasound.
- It's caused by the interference of the returning wave collected by the transducer.

- It can be expressed by:

$$g(n, m) = f(n, m) * u(n, m) + \xi(n, m)$$

- where $g(n, m)$ is the observed image, $u(n, m)$ multiplicative component and $\xi(n, m)$ additive component of the speckle noise.

Image noise

Poisson noise (or Shot noise):

- Mainly appears in the brighter parts of the image and is caused by fluctuations of the number of photons detected.
- Probability distribution function for each pixel is given by:

$$p(f = k) = \frac{k^\lambda e^{-\lambda}}{k!}$$

- $\lambda > 0$ is a Poisson distribution parameter.

Image noise



(Original)



a. Salt & pepper noise



b. Gaussian noise



c. Speckle noise



d. Poisson noise

Various image noise types [SAR2016].

Bibliography

- [PIT2021] I. Pitas, “Computer vision”, Createspace/Amazon, in press.
- [PIT2017] I. Pitas, “Digital video processing and analysis” , China Machine Press, 2017 (in Chinese).
- [PIT2013] I. Pitas, “Digital Video and Television” , Createspace/Amazon, 2013.
- [NIK2000] N. Nikolaidis and I. Pitas, “3D Image Processing Algorithms”, J. Wiley, 2000.
- [PIT2000] I. Pitas, “Digital Image Processing Algorithms and Applications”, J. Wiley, 2000.
- [MIS2017] Mishra VK, Kumar S, Shukla N. Image acquisition and techniques to perform image acquisition. SAMRIDDHI: A Journal of Physical Sciences, Engineering and Technology. 2017 Jun 25;9(01):21-4.
- [LIT2001] Litwiller D. Ccd vs. cmos. Photonics spectra. 2001 Jan;35(1):154-8.
- [EDM2021] Edmund Optics, “Understanding Camera Sensors for Machine Vision Applications”, <https://www.edmundoptics.eu/knowledge-center/application-notes/imaging/understanding-camera-sensors-for-machine-vision-applications/>

Bibliography

- [COP2021] CopyFaxes, “The Scanning Process”, <https://copyfaxes.com/blog/scanning-process/>
- [REA2021] ReadyArtwork , “Scalable Vector Graphic Examples”, <https://www.readyartwork.com/category/design/page/3/>
- [BAY2021] “Bayer Filter”, https://en.wikipedia.org/wiki/Bayer_filter
- [RAH2021] S. F. Rahman, “CSS Techniques for Retina Displays”, <https://www.sitepoint.com/css-techniques-for-retina-displays/>
- [WIL2021] B. Williams, “How To Use White Balance In Photography”, <https://bwillcreative.com/how-to-use-white-balance-in-photography/>
- [ROS2021] A. Rosebrock, “OpenCV Gamma Correction”, <https://www.pyimagesearch.com/2015/10/05/opencv-gamma-correction/>
- [WOO2007] Woodman, R., 2007. A photometric stereo approach to face recognition.
- [PHO2013] Photo Review Issue 57, 2013
- [SAR2016] Sarker O, Akter S, Mishu AA. Review on the performance of different types of filter in the presence of various noises. Engineering International. 2016 Dec 31;4(2):49-56.

Q & A

Thank you very much for your attention!

**More material in
<http://icarus.csd.auth.gr/cvml-web-lecture-series/>**

**Contact: Prof. I. Pitas
pitass@csd.auth.gr**