Topics on Robotics & Vision

B.Sc

Digital Image Processing

LectureSlide

D. T. McGuiness, Ph.D

WS.2025

Table of Contents

Introduction

Introduction

- The goal of this lecture is to introduce you to image processing and its wide applications in industry.
- We shall have a wide focus on the technologies and methods which make image processing an essential discipline for engineers.
- This lecture is a total of 4 SWS with a total of sixty (60) UE.
- A unit (UE) is defined as 45 min lecture.

- Lecture materials and all possible supplements will be present in its Github Repo.
  - You can easily access the link to the web-page from here.

> Github is chosen for easy access to material management and CI/CD capabilities and allowing hosting websites.

- In the lecture content is also distributed as a WebBook which can be accessed from the Repo website.

- The student should be comfortable with working with physical problems and have a basic understanding of material science along with calculus.

Requirements	Taught Lecture	Code	Degree	Outcome
Python	Programming I	PRG I	B.Sc	Python Programming
Linear Algebra	Mathematics I	MAT I	B.Sc	Signal Processing
				Image Processing
-				Camera Technology
-				Statistical Analysis

Table 1: Distribution of materials across the semester.

Description	Value
Official Name	Image Processing
Lecture Code	IMP
Module Code	MECH-B-5-MRV-IMP-ILV
Lecture Name	Digital Image Processing
Semester	5
Season	WS
Lecturer	Daniel T. McGuiness, Ph.D
Module Responsible	BnM
Software	Python
SWS Total	4
UE Total	60
ECTS	5
Working Language	English

- The lecture will have a single personal assignment comprising of a set list of questions which you can use programming languages to solve on your own.
- There will also be a group assignment where you will team up with your classmates to come up with ideas for applying image processing concepts to problems.

Assignment Type	Value
Personal Assignment	40
Group Project	60
Sum	100

Title
Fundamentals of Image Processing
Computer Vision: Algorithms and Applications
Feature Extraction and Image Processing for Computer Vision
Digital Image Processing
Types Of Camera Sensor
Introduction To Quantum Efficiency
Dark Current
Linearity - Imaging Topics

Table 2: Lecture sources which can be useful during the course of the lecture. For more information on sources, please consult the repo.

Topic	Units	Self Study
Mathematical Fundamentals	4	8
Perception	4	8
Image Formats	4	8
Camera	4	8
Display	4	8
Noise	4	8
Histogram Operations	4	8
Morphological Opeations	4	8
Blurring Filters	4	8

Introduction ii

Topic	Units	Self Study
Feature Analysis	4	8
Edge Detection	4	8
Introduction to Artificial Neural Networks	4	8
Computer Vision using Convolutional Neural Networks	4	8
SUM	52	104

Mathematical Fundamentals

(LO1) An Overview of mathematical methods,

(LO2) A revisit on convolution.

(LO3) Definitions of analogue and digital,

(LO4) A mathematical look into Discrete Fourier Transform (DFT).

- Computer Vision encompasses multiple disciplines, including digital image processing, cameras, displays, filters, and transform.
- To better prepare, it is important to refresh/learn some mathematical principles and concepts.

Concepts and Principles

- Principle of convolution,
- Discrete Fourier analysis,
- Shannon-Nyquist Sampling Theorem,
- A brief introduction to Information Theory,
- The concept of information entropy.

- Convolution, mathematically is defined as:

$$(f * g)(t) = \int_{-\infty}^{+\infty} f(\tau)\, g(t-\tau)\ \mathrm{d}\tau.$$

- Where f and g are arbitrary function and $*$ is the convolution operator.

> To put it simply, convolution is just fancy multiplication. But in principle, it has a strong relationship with Laplace transform and is used significantly in image processing.

Example

Imagine you manage a hospital treating patients with a single disease.

You have:

Treatment Plan 3 Every patient gets 3 units of the cure on their first day.

Patient List `[1, 2, 3, 4, 5]` Your patient count for the week (1 person Monday, 2 people on Tuesday, etc.).

How much medicine do you use each day?

Solution

The answer is a quick multiplication:

$$\text{Plan} \times \text{Patients} = \text{Daily Usage}$$
$$3 \times [1, 2, 3, 4, 5] = [3, 6, 9, 12, 15]$$

Multiplying the plan by the patient list gives usage for upcoming days:

$$[3, 6, 9, 12, 15]$$

Everyday multiplication of (3×4) means using the plan with a single day of patients:

$$[3] \times [4] = [12]$$

Solution

Now the disease mutates and needs multi-day treatment. A new plan:

$$\text{Plan:}[3, 2, 1]$$

Meaning:

- 3 units of the cure on day one,
- 2 units on day two,
- 1 unit on day three.

Given the same patient schedule of:

$$\text{Patient:}[1, 2, 3, 4, 5]$$

what's our medicine usage each day? Let's see

Solution

- On day 1, 1 patient A comes in. It's their first day, so 3 units.
- On day 2, A gets 2 units (second day), but two new patients (B1 & B2) arrive, who get 3 each ($2 \times 3 = 6$).
  - The total is $2 + (2 \times 3) = 8$ units.
- On Wednesday, it's trickier: The patient A finishes (1 unit, her last day), the B1 and B2 get 2 units (2 * 2), and there are 3 new Wednesday people ...

The patients are overlapping and it's hard to track. How can we organise this calculation?

Solution

An idea worth considering is to reverse the order of the patient list:

$$\text{New Patient List:}[5, 4, 3, 2, 1]$$

Next, imagine we have 3 separate rooms where we apply the proper dose:

$$\text{Rooms:}[3, 2, 1]$$

On your first day, you walk into the first room and get 3 units of medicine. The next day, you walk into room #2 and get 2 units. On the last day, you walk into room #3 and get 1 unit. There's no rooms afterwards, and your treatment is done.

Solution

To calculate the total medicine usage, line up the patients and walk them through the rooms:

C.R. 1

text

```
Monday
----------------------------
Rooms                  3 2 1
Patients       5 4 3 2 1

Usage                  3
```

On Monday (our first day), we have a single patient in the first room. A gets 3 units, for a total usage of 3.

Makes sense, right?

Solution

On Tuesday, everyone takes a step forward:

C.R. 2

text

```
Tuesday
----------------------------
Rooms                      3 2 1
Patients ->          5 4 3 2 1

Usage                      6 2      = 8
```

The first patient is now in the second room, and there's 2 new patients in the first room. We multiply each room's dose by the patient count, then combine.

Solution

C.R. 3

text

```
Wednesday
----------------------------
Rooms                  3 2 1
Patients ->        5 4 3 2 1
Usage                  9 4 1    = 14

Thursday
-----------------------------
Rooms                  3 2 1
Patients ->          5 4 3 2 1
Usage                 12 6 2    = 20

Friday
-----------------------------
Rooms                  3 2 1
Patients ->            5 4 3 2 1
Usage                 15 8 3    = 26
```

Solution

It's intricate, but we figured it out, right? We can find the usage for any day by reversing the list, sliding it to the desired day, and combining the doses.

The total day-by-day usage looks like this (don't forget Sat and Sun, since some patients began on Friday):

C.R. 4

```python
Plan      *  Patient List   = Total Daily Usage

[3 2 1]   *  [1 2 3 4 5]    = [3 8 14 20 26 14 5]
              M T W T F        M T W  T  F  S  S
```

This calculation is the convolution of the plan and patient list. It's a fancy multiplication between a list of input numbers and a "program".

Example

Write a script which does convolution of the following two (2) arrays:

$$A = [1,\ 1,\ 2,\ 2,\ 1] \qquad\qquad B = [1,\ 1,\ 1,\ 3]$$

Solution

C.R. 5

python

```python
import numpy as np
def convolve_1d(signal, kernel):
    kernel = kernel[::-1]
    k = len(kernel)
    s = len(signal)
    signal = [0]*(k-1)+signal+[0]*(k-1)
    n = s+(k-1)
    res = []
    for i in range(s+k-1):
        res.append(np.dot(signal[i:(i+k)], kernel))
    return res
```

Solution

C.R. 6

python

```python
A = [1,1,2,2,1]
B = [1,1,1,3]

print(convolve_1d(A, B))
```

- An operation on two functions (f and g) that produces $f * g$.

It expresses how the shape of one is modified by the other.

- There are several notations to indicate convolution with the most common is:

$$c = f(t) * g(t) = (f * g)(t),$$

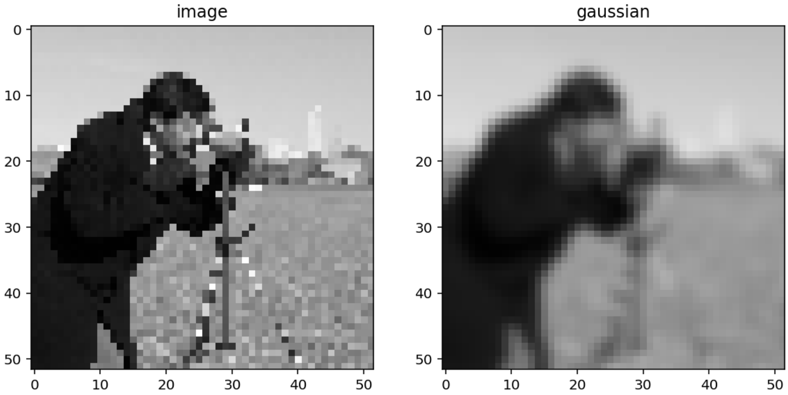


Figure 1: An example of convolution used in image processing. Here a pixellated image is smoothened out using Gaussian Blur which relies on convolution.

- In 2D continuous space (i.e., analogue):

$$\begin{aligned} c(x, y) &= f(x, y) * g(x, y), \\ &= \int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(\chi, \xi)\, g(x - \chi, y - \xi)\, \mathrm{d}\chi\, \mathrm{d}\xi. \end{aligned}$$

- In 2D discrete space (i.e., digital):

$$\begin{aligned} c[m, n] &= f[m, n] * g[m, n], \\ &= \sum_{j=-\infty}^{+\infty} \sum_{k=-\infty}^{+\infty} f[j, k]\, g[m - j, n - k]. \end{aligned}$$

- It is the single most important technique in digital signal processing.
- Using the strategy of impulse decomposition, systems are described by a signal called the impulse response.
- Convolution is important as it relates the three (3) signals of interest:
  1. Input signal,
  2. Output signal,
  3. Impulse response.
- But now, let's look at some of it's properties:

Commutative

- The order in which we convolve two signals does **NOT** change the result:

$$f(t) * g(t) = g(t) * f(t)$$

Distributive

- if there are three signals $f(t), g(t), h(t)$, then the convolution of $f(t)$ is said to be distributive:

$$f(t) * [g(t) + h(t)] = [f(t) * g(t)] + [f(t) * h(t)]$$

Associative

- The way in which the signals are grouped in a convolution does not change the result:

$$f(t) * [g(t) * h(t)] = [f(t) * g(t)] * h(t)$$

Shift Property

- The convolution of a signal with a time shifted signal results a shifted version of that signal. i.e.,

$$f(t) * g(t) = y(t)$$

- Then according to the shift property of convolution:

$$f(t) * f(t - T_0) = y(t - T_0)$$

- Similarly:

$$f(t - T_0) * f(t) = y(t - T_0)$$

- Therefore:

$$f(t - T_1) * f(t - T_2) = y(t - T_1 - T_2)$$

Mathematical Fundamentals

$$
\underbrace{\begin{pmatrix}
0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 1 & 1 & 1 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 \\
0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 \\
0 & 1 & 1 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}}_{I}
*
\underbrace{\begin{pmatrix}
1 & 0 & 1 \\
0 & 1 & 0 \\
1 & 0 & 1
\end{pmatrix}}_{K}
=
\underbrace{\begin{pmatrix}
1 & 4 & 3 & 4 & 1 \\
1 & 2 & 4 & 3 & 3 \\
1 & 2 & 3 & 4 & 1 \\
1 & 3 & 3 & 1 & 1 \\
3 & 3 & 1 & 1 & 0
\end{pmatrix}}_{I * K}
$$

Figure 2: A visual representation of how convolution works in 2D.

- Converting from a continuous 2D data $a\,(x,\, y)$ to its digital representation $a\,[x,\, y]$ requires the process of sampling.
- An ideal sampling system is defined as the image $a\,(x,\, y)$ multiplied by an ideal 2D impulse train $\delta\,(x,\, y)$:

$$b\,[m,\, n] = a\,(x,\, y) \sum_{m=-\infty}^{+\infty} \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} \delta\,(x - mX_0,\, y - nY_0)$$
$$= \sum_{m=-\infty}^{+\infty} \sum_{n=-\infty}^{+\infty} a\,(mX_0,\, nY_0)\,\delta\,(x - mX_0,\, y - nY_0)\,.$$

  where X_0 and Y_0 are the sampling distance or intervals and δ is the Dirac delta function.
- If you were to sample in square shapes $X_0 = Y_0$ where you could think of each individual block a pixel

- To reconstruct a continuous analog signal from its sampled version accurately, the sampling rate must be at least twice the highest frequency present in the signal.
- This ensures that there are enough samples taken per unit of time to capture all the details of the original waveform without introducing aliasing, which can cause distortion or artifacts in the reconstructed signal.

$$f_s \geq 2f_m$$

where f_s is the signal frequency, f_m is the maximum sample frequency.

> This is only a theoretical limit and **NOT** a practical one.

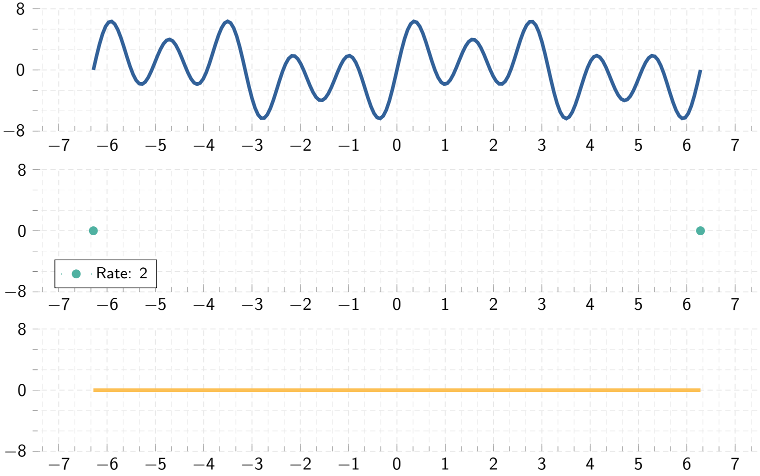


Figure 3: The effects of signal reconstruction on the sampling rate.

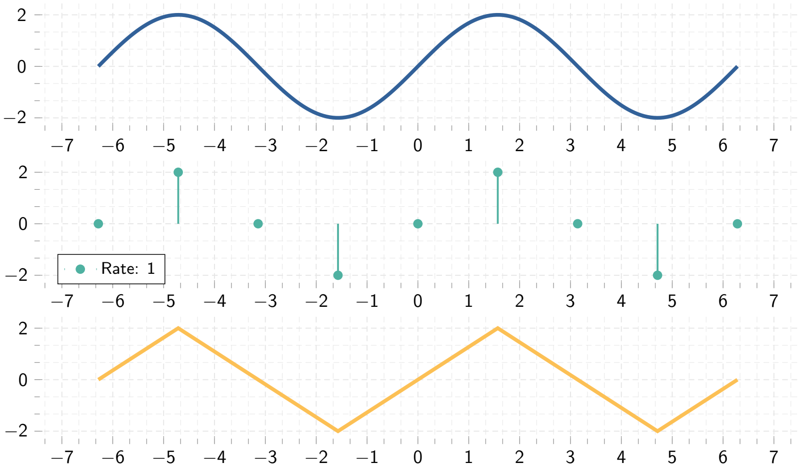


Figure 4: Reconstruction of the signal with 1 times the signal frequency.

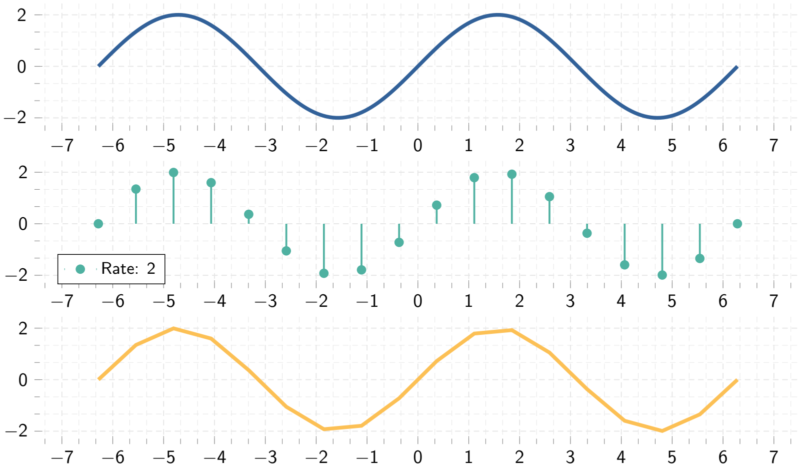


Figure 5: Reconstruction of the signal with 2 times the signal frequency.

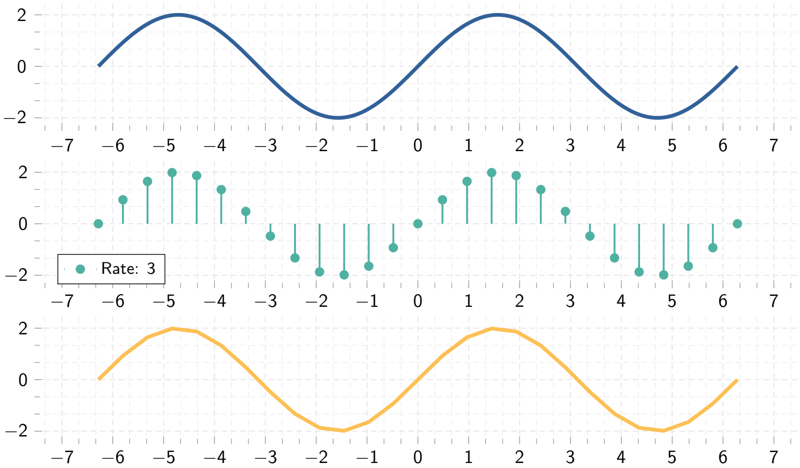


Figure 6: Reconstruction of the signal with 3 times the signal frequency.

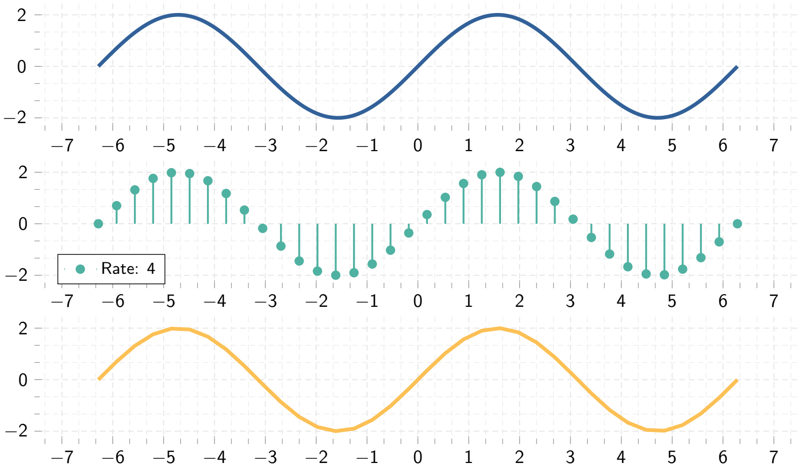


Figure 7: Reconstruction of the signal with 4 times the signal frequency.

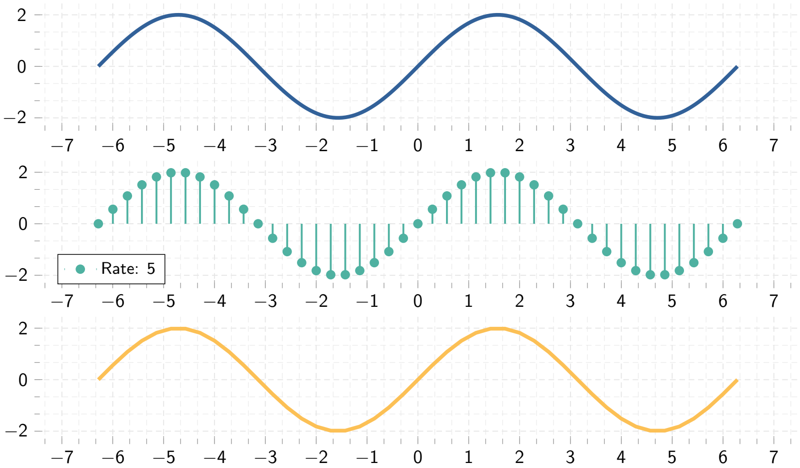


Figure 8: Reconstruction of the signal with 5 times the signal frequency.

- In practice, doubling frequency is **NOT** enough recreate the signal.
- Approaching *Nyquist frequency* will create a siren like sound, and reaching exact frequency will record a pulse-wave approximation of a sine wave at an amplitude which will vary based on phase.
- Even 4 times sampling will only reconstruct a triangle wave and shifting the phase will create tonal distortion.

For practical cases at least 6 times sampling rate is needed to accurately reconstruct the sine wave.

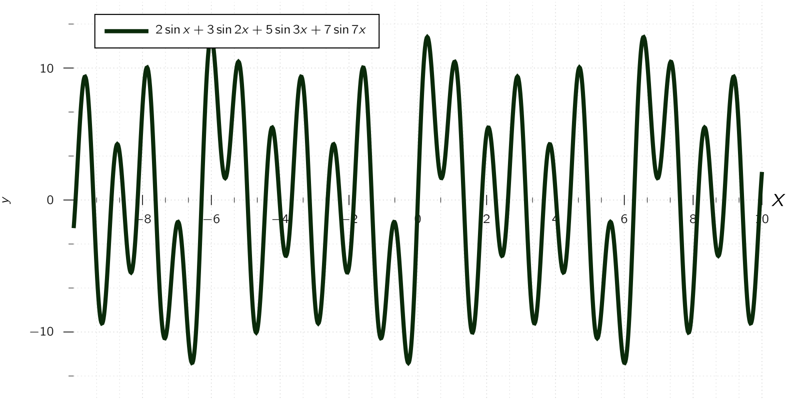


Figure 9: A sample signal with containing sample sine waves.

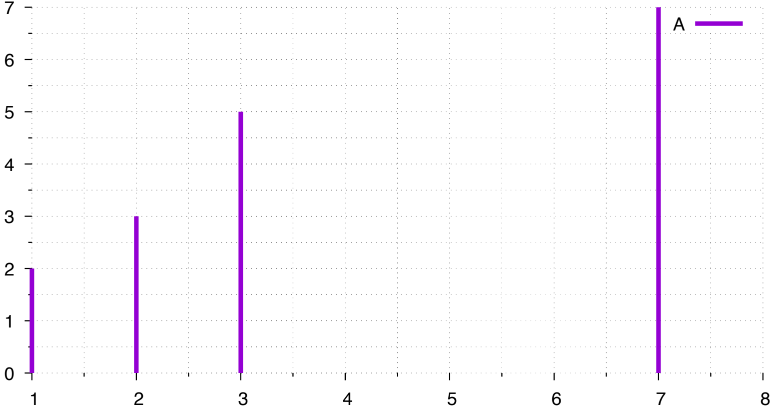


Figure 10: The FFT of the previous complex signal.

- Two (2) key problems arise when conducting spectral analysis of finite, discrete time series (not an infinite time series):

Aliasing Only resolving frequencies lower than the *Nyquist frequency* and higher frequencies get aliased to lower frequencies.

Spectral Leakage we assume all wave-forms stop and start at 0 and end at n, but in the real world, many wave-numbers may not complete a full integer number of cycles throughout the domain, causing spectral leakage to other wave numbers.

- If the initial samples are **NOT** sufficiently closely spaced to represent high-frequency components present in the underlying function, then the DFT values will be corrupted by aliasing.
- The solution is either to increase the sampling rate (if possible) or to pre-filter the signal in order to minimise its high-frequency spectral content.

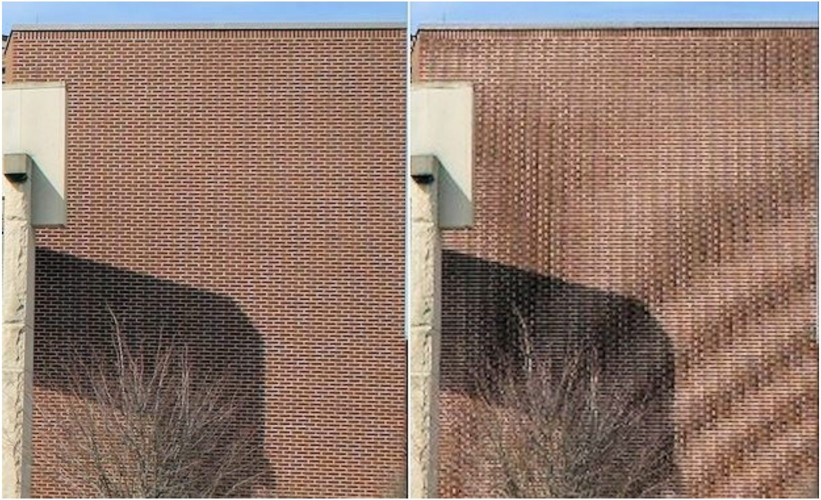


Figure 11: An example of under-sampling an image. Here aliasing produces non-real distortions of digitized images.

- The continuous Fourier transform of a periodic waveform requires the integration to be performed over the interval $-\infty$ to $+\infty$ or over an integer number of cycles of the waveform.
- If we attempt to complete the DFT over a non-integer number of cycles of the input signal, might cause the transform to be corrupted in some way.
- Let's start with looking at a simple sine-wave

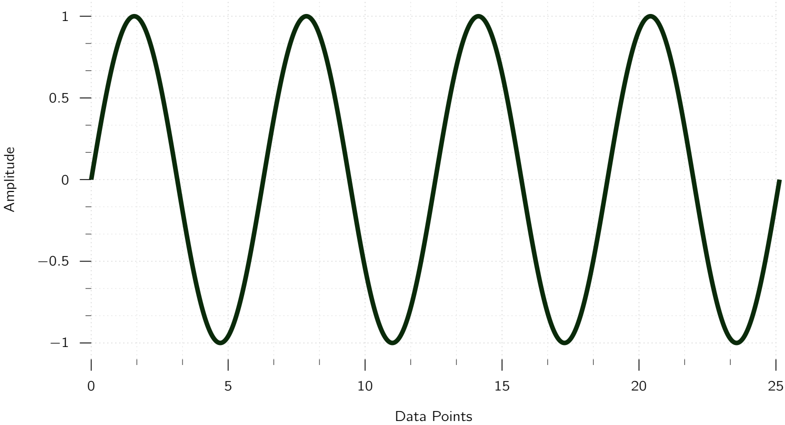


Figure 12: An example of a sine wave with four (4) complete cycles.

- Computing the discrete power spectrum gives:

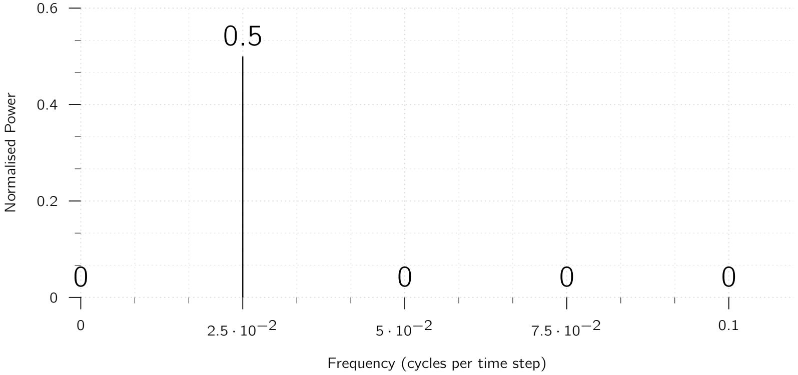


Figure 13: The PSD of an un-windowed sine wave.

- As expected, a **single spectral peak** corresponding to the frequency of our sine wave.

- Let's see what happens if we apply a window to our sine wave which **cuts off** the sine wave such that the sine function does not complete an integer number of cycles within the time domain.

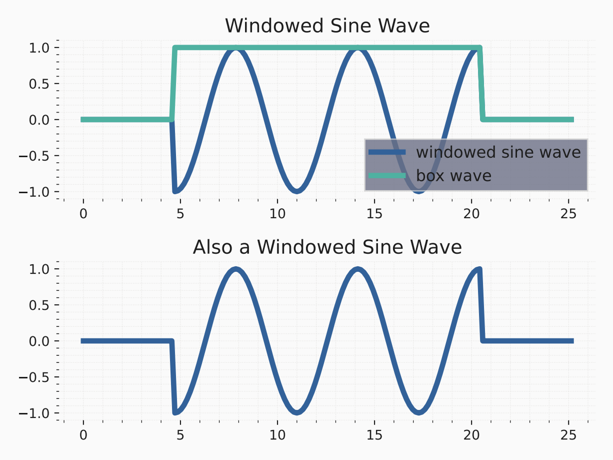


Figure 14: Windowed sine wave.

- To demonstrate spectral leakage, we will now compute the discrete power spectrum of the windowed sine wave to see what happens.

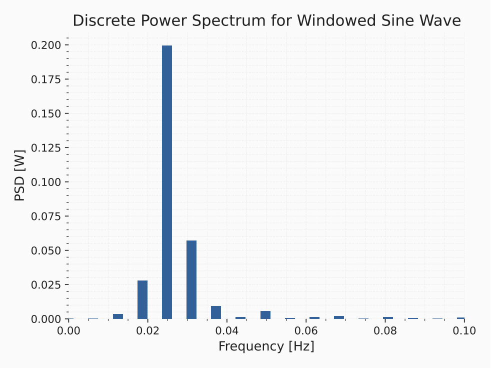


Figure 15: PSD of a windowed sine wave.

Example

Below is a signal with 1 Hz, Amplitude of 1 and 8 Sampling points.

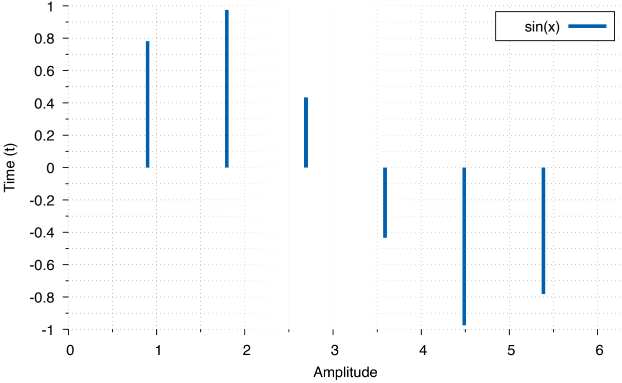


Figure 16: A Sampled Sine wave.

Solution

As it is a single sine function with 1 Hz, we expect a single value of 1 in the frequency domain (at 1 Hz).

The sampling points will sample the signal and retrieve the following data points as shown in the array below:

$$x_k = \begin{bmatrix} 0 & 0.707 & 1 & 0.707 & 0 & -0.707 & -1 & -0.707 \end{bmatrix}$$

Solution

Once we have these sampling points (x_n), we can turn our attention to the DFT formula:

$$X_k = \sum_{n=0}^{N-1} x_n \cdot e^{-(\mathbf{j}\,2\pi kn)/\ N}$$

where X_k is the k^{th} frequency bin.

For $x_0 = 0$ the exponential is removed and are left with $X_0 = 0$.

Solution

For the cases of X_1:

$$X_1 = \sum_{n=0}^{7} x_\text{n} \cdot e^{-(\mathbf{j}\, 2\pi\,(1)\, n)/\ N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^\mathsf{T} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\,2\pi\,(1\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(2\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(3\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(4\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(5\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(6\times 1))/\ N} \\ e^{-(\mathbf{j}\,2\pi\,(7\times 1))/\ N} \end{bmatrix} = 0 - \mathbf{j}\,4 \quad \blacksquare$$

Solution

For the cases of X_2:

$$X_2 = \sum_{n=0}^{7} x_n \cdot e^{-(\mathbf{j}\, 2\pi\, (2)\; n)/\; N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathsf{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\, 2\pi\, (1 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (2 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (3 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (4 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (5 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (6 \times 2))/\; N} \\ e^{-(\mathbf{j}\, 2\pi\, (7 \times 2))/\; N} \end{bmatrix} = 0 \quad \blacksquare$$

Solution

For the cases of X_3:

$$X_3 = \sum_{n=0}^{7} x_n \cdot e^{-(\mathbf{j}\,2\pi\,(3)\,n)/\,N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathsf{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\,2\pi\,(1\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(2\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(3\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(4\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(5\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(6\times 3))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(7\times 3))/\,N} \end{bmatrix} = 0 \quad \blacksquare$$

Solution

For the cases of X_4:

$$X_4 = \sum_{n=0}^{7} x_{\mathrm{n}} \cdot e^{-(\mathbf{j}\, 2\pi\, (4)\, n)/\ N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathsf{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\, 2\pi\, (1 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (2 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (3 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (4 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (5 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (6 \times 4))/\ N} \\ e^{-(\mathbf{j}\, 2\pi\, (7 \times 4))/\ N} \end{bmatrix} = 0 \quad \blacksquare$$

Solution

For the cases of X_5:

$$X_5 = \sum_{n=0}^{7} x_n \cdot e^{-(\mathbf{j}\,2\pi\,(5)\,n)/\,N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\,2\pi\,(1\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(2\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(3\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(4\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(5\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(6\times5))/\,N} \\ e^{-(\mathbf{j}\,2\pi\,(7\times5))/\,N} \end{bmatrix} = 0 \quad \blacksquare$$

Solution

For the cases of X_6:

$$X_6 = \sum_{n=0}^{7} x_{\mathrm{n}} \cdot e^{-(\mathbf{j}\,2\pi\,(6)\;n)/\;N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathrm{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\,2\pi\,(1\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(2\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(3\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(4\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(5\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(6\times 6))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(7\times 6))/\;N} \end{bmatrix} = 0 \quad \blacksquare$$

Solution

For the cases of X_7:

$$X_7 = \sum_{n=0}^{7} x_{\text{n}} \cdot e^{-(\mathbf{j}\,2\pi\,(7)\;n)/\;N}$$

$$= \begin{bmatrix} 0 \\ 0.707 \\ 1 \\ 0.707 \\ 0 \\ -0.707 \\ -1 \\ -0.707 \end{bmatrix}^{\mathsf{T}} \cdot \begin{bmatrix} 0 \\ e^{-(\mathbf{j}\,2\pi\,(1\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(2\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(3\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(4\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(5\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(6\times 7))/\;N} \\ e^{-(\mathbf{j}\,2\pi\,(7\times 7))/\;N} \end{bmatrix} = 0 + \mathbf{j}\,4 \quad \blacksquare$$

Solution

- Therefore the values are of the transform are:

$$X_k = \begin{bmatrix} 0 & 0 - \mathbf{j}4 & 0 & 0 & 0 & 0 & 0 & 0 + \mathbf{j}4 \end{bmatrix}$$

- We can see only the first and the seventh bins have values other than zero.
- Calculating the magnitudes of the bins, we arrive at 4.

$$|X_k| = \begin{bmatrix} 0 & 4 & 0 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}$$

Solution

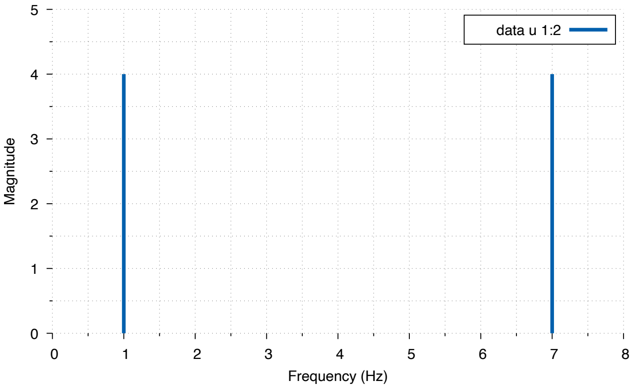


Figure 17: Sampled dataset of the original signal. There is still another step.

Solution

- The frequency resolution of the plot is the sampling frequency divided by the number of samples:

$$\text{Resolution} = \frac{\text{Sampling Frequency}}{\text{Number of Samples}}$$

- This means we can get values for every integer frequency values.

Solution

- We can see we get a value for the first frequency bin (1 Hz) and it makes sense.
- The reason we get a frequency bin is due to the plot being a two-sided frequency plot where it shows the energy in both the positive and negative frequency.

> The negative frequencies are always complex conjugate to the positive frequencies, so there is no additional information in the negative frequencies.

Solution

- Therefore, to convert from a two-sided spectrum to a single-sided spectrum, we discard the second half of the array and multiply every point except for DC by two (2).
- The last operation is to divide the magnitudes of the lower frequencies by the number of samples used in deriving these bins:

$$\boldsymbol{X}_{\mathrm{k}} = \begin{bmatrix} 0 & 8 & 0 & 0 \end{bmatrix}$$

$$\boldsymbol{X}_{\mathrm{k}}/N = \begin{bmatrix} 0 & 1 & 0 & 0 \end{bmatrix} \quad \blacksquare$$

- The sum (or integral) of the square of a function is equal to the sum (or integral) of the square of its transform.
- For continuous signals:

$$\int_{-\infty}^{\infty} \left|f\left(t\right)\right|^2 \,\mathrm{d}t = \frac{1}{2\pi}\int_{-\infty}^{\infty} \left|F\left(\omega\right)\right|^2 \,\mathrm{d}\omega = \int_{-\infty}^{\infty} \left|F\left(2\pi f\right)\right|^2$$

This signal energy is **NOT** to be confused with physical energy.

Average Value (μ)

- Defined as the sample mean of a given region.
- The equation is defined as below (it is also known as expected value):

$$\mu = \frac{1}{n} \sum_{i=0}^{n} x_{\mathrm{n}}$$

Standard Deviation (σ)

- The standard deviation is a measure of the amount of variation of the values of a variable about its mean (μ):

$$\sigma = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (x_i - \mu)^2}$$

Mode

- The mode is the value appears most often in a set of data values. i.e., in an data pool of:

$$X = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 2 & 7 & 2 & 9 \end{bmatrix}$$

- The mode of is 2 as it is the most frequent value of the data set. whereas in:

$$X = \begin{bmatrix} 2 & 4 & 9 & 6 & 4 & 6 & 6 & 2 & 8 & 2 \end{bmatrix}$$

the mode is (2, 6) as there are two (2) values with same frequency.

Median

- The median is the middle value separating the greater and lesser halves of the data set.
- For a ordered data set X with n elements,
  - if n is odd:

$$\operatorname{med}(x) = x\frac{n+1}{2},$$

  - if n is even,

$$\operatorname{med}(x) = x\left(\frac{n}{2}\right) + x\frac{\frac{n}{2}+1}{2}$$

- i.e., in a ordered data set of:

$$X = \begin{bmatrix} 1 & 2 & 2 & 3 & 4 & 7 & 9 \end{bmatrix},$$

- the median is 3 and in:

$$Y = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 8 & 9 \end{bmatrix}.$$

the median is 4.5.

- The signal-to-noise ratio (SNR) can have several definitions depending on the field.
- Noise is characterised by its standard deviation, σ.
- The characterisation of the signal can differ.

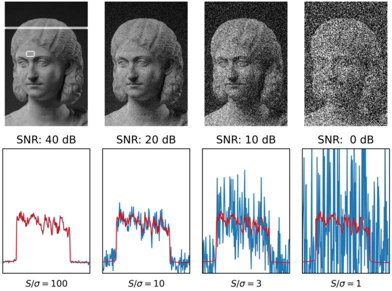


Figure 18: A gray-scale photography with different signal-to-noise ratios (SNRs).

- If the signal is known to lie between two (2) boundaries:

$$a_{\text{min}} \leq a \leq a_{\text{max}}$$

then the SNR is defined as:

$$\text{SNR} = 20 \log_{10} \left(\frac{a_{\text{max}} - a_{\text{min}}}{s_{\text{n}}} \right) \text{ dB.}$$

- If the signal is not bounded but has a statistical distribution then two other definitions are known:

$$\text{SNR} = 20 \log_{10} \left(\frac{\mu}{\sigma} \right) \text{ dB.}$$

- In 1948, Claude Shannon published a paper called A Mathematical Theory of Communication.
- This paper heralded a transformation in our understanding of information.
- Before Shannon's paper, information had been viewed as a kind of poorly defined ethereal concept.
- But after Shannon's paper, it became apparent that information is a well-defined and, above all, measurable quantity.

- Information theory defines definite, unbreachable limits on precisely how much information can be communicated between any two (2) components of any system,
  - whether this system is **man-made** or **natural**.
- The basic laws of information can be summarised as follows.

1. there is a upper limit, the channel capacity, to the amount of information that can be communicated through that channel,
2. this limit shrinks the amount of noise in the channel increases,

This limit can be approached by clever methods of encoding data.

- The word bit is derived from binary digit,
  - but a bit and a binary digit are fundamentally different types of quantities.

A binary digit is the value of a binary variable, whereas a bit is an amount of information.

- Consider a coin which lands heads up 90% of the time:

$$p\left(x_{\mathrm{h}}\right) = 0.9.$$

- When this coin is flipped, we expect it to land heads up ($x = x_{\mathrm{h}}$),
- When it does, we are less surprised than when it lands tails ($x = x_{\mathrm{t}}$).

> The more improbable a particular outcome is, the more surprised we are to observe it.

- If we use $\log_2$ then the Shannon information or surprisal of each outcome is measured in bits.

$$\text{Shannon Information} = \log_2 \frac{1}{p\left(x_{\mathrm{h}}\right)}$$

Mathematical Fundamentals

- We can represent the coin flip outcome as random variable x,
  - such that a head is $x = x_{\mathrm{h}}$ and a tail is $x = x_{\mathrm{t}}$.
- In practice, we are not usually interested in the surprise of a particular value of a random variable, but we are interested in how much surprise, on average, is associated with the entire set of possible values.
- The average surprise of a variable x is defined by its probability distribution $p(x)$, and is called the entropy of $p(x)$, represented as $H(x)$.

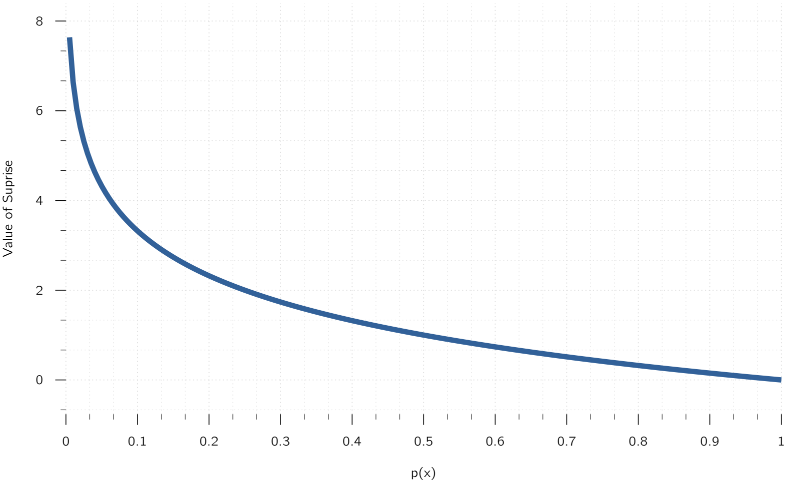


Figure 19: The quantifiable surprise with respect to increasing probability.

- If a coin is fair or unbiased then:

$$p_{x_h} = p_{x_t} = 0.5$$

- The Shannon information gained when a head or a tail is observed is:

$$\log 1/0.5 = 1 \text{ bit}$$

- The average Shannon information gained after each coin flip is also 1 bit.
- Because entropy is defined as average Shannon information, the entropy of a fair coin is $H(x) = 1$ bit.

- Let's look at a biased coin with a probability of a head is $p(x_{\mathrm{h}}) = 0.9$.
  - it is easy to predict the result of each coin flip (i.e. with 90% accuracy if we predict a head for each flip)
- If the outcome is a head then the amount of Shannon information gained is

$$log(1/0.9) = 0.15 bits.$$

- But if the outcome is a tail then the amount of Shannon information gained is:

$$log(1/0.1) = 3.32 bits.$$

- Notice that more information is associated with the more surprising outcome.

Entropy of an Unfair Coin

- Given that the proportion of flips that yield a head is p(xh), and that the proportion of flips that yield a tail is p(xt) (where p(xh) + p(xt) = XX), the average surprise is

$$H(x) = p(x_{\mathrm{h}}) \log \frac{1}{p(x_{\mathrm{h}})} + p(x_{\mathrm{t}}) \log \frac{1}{p(x_{\mathrm{t}})},$$

- Which comes to XX bits.
- If we define a tail as $x_1 = x_{\mathrm{t}}$ and a head as $x_2 = x_{\mathrm{h}}$ then the above equation is written as:

$$H(x) = \sum_{i=1}^{2} p(x_{\mathrm{i}}) \log \frac{1}{p(x_{\mathrm{i}})} \text{ bits.}$$

- More generally, a random variable x with a probability distribution

$$p(x) = p(x_1), \ldots, p(x_m)$$

has an entropy of

$$H(x) = \sum_{i=1}^{m} p(x_i) \log \frac{1}{p(x_i)} \text{ bits.}$$

- Entropy is a measure of uncertainty.
- When our uncertainty is reduced, we gain information,
  - so information and entropy are two sides of the same coin.
- However, information has a rather subtle interpretation, which can easily lead to confusion.
- Average information shares the same definition as entropy,
  - but whether we call a given quantity information or entropy depends on whether it is being given to us or taken away.

- For example, if a variable has high entropy the initial uncertainty of the variable is large and is, by definition, exactly equal to its entropy.
- If we are told the variable value, on average, we have been given information equal to the uncertainty (entropy) we had about its value.
- Thus, receiving an amount of information is equivalent to having exactly the same amount of entropy (uncertainty) taken away.

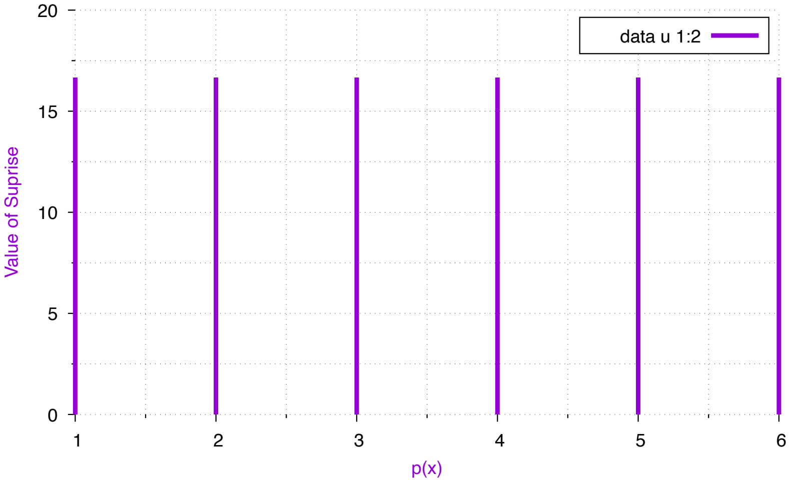


Figure 20: The probability distribution of 1 dice(s).

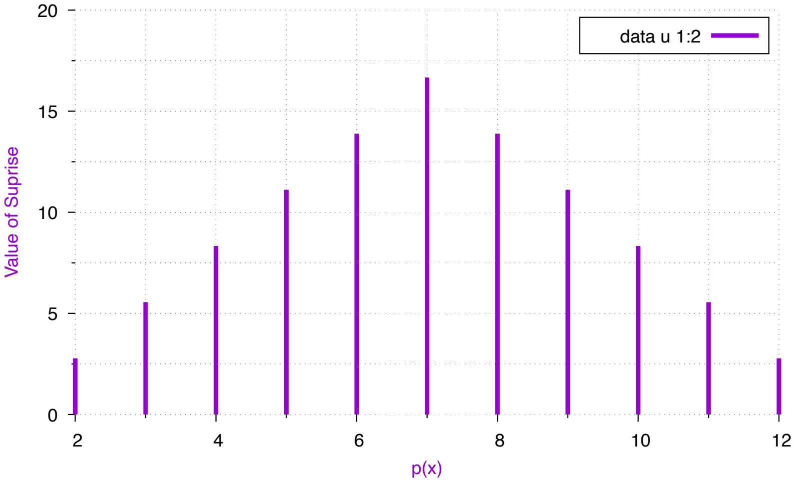


Figure 21: The probability distribution of 2 dice(s).

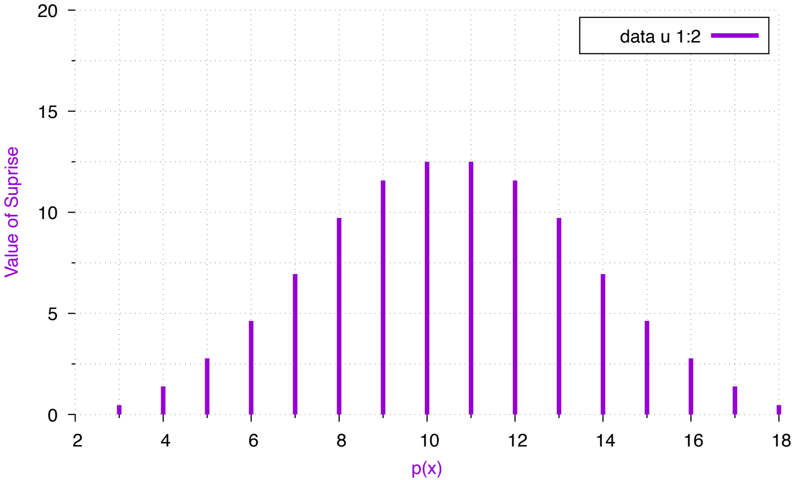


Figure 22: The probability distribution of 3 dice(s).

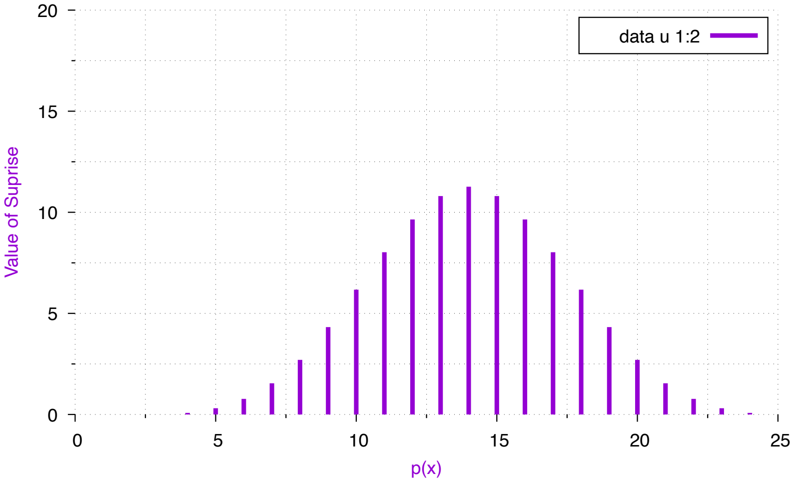


Figure 23: The probability distribution of 4 dice(s).

- Throwing a pair of 6-sided dice produces an outcome in the form of an ordered pair of numbers.
    - There are a total of 36 equiprobable outcomes,
- If we define an outcome value as the sum of this pair of numbers then there are $m = 11$ possible outcome values:

$$A_x = \{2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12\} .$$

Dividing the frequency of each outcome value by 36 yields the probability p of each outcome value.

- We can use these 11 probabilities to find the entropy.

$$
\begin{aligned}
H(x) &= p(x_1)\log\frac{1}{p(x_1)} + p(x_1)\log\frac{1}{p(x_1)} + \cdots + p(x_{11})\log\frac{1}{p(x_{11})} \\
&= 3.27 \text{ bits.}
\end{aligned}
$$

Perception

(LO1) A Look into Human Vision,

(LO2) Definition of Colour and Standardisation,

(LO3) How Vision is Perceived,

(LO4) Types of Colour-spaces.

- Many image processing applications are intended to produce images to be viewed by humans.
    - This is in contrast to industrial robots.
- It is important to understand the characteristics and limitations of the human visual system [1].
- At the outset it is important to realise:
    1. The human visual system is not well understood [1].
        - It is not easy to study the human visual system without directly measuring it.
    2. No objective measure exists for judging the quality of an image that corresponds to human assessment of image quality,
        - A colour you find fitting might be repugnant to someone.
    3. A typical human observer does not exist.

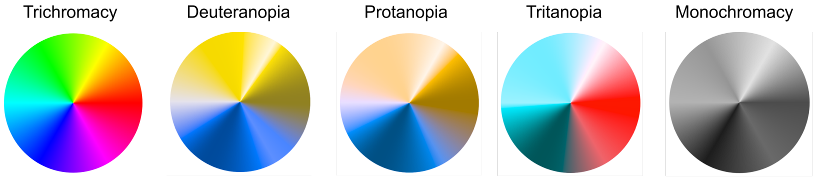


Figure 24: A color wheel depicted approximately as it would be seen by a person with different kinds of color vision or color blindness: trichromacy (normal colour vision), deuteranopia (red-green color blind), protanopia (red-green colour blind), tritanopia (blue-yellow color blind), and monochromacy (completely colour blind). [2].

- Statistically 1 in 8 men and 1 in 200 women have a form of colour blindness [3].
- Men have one ***X*** chromosome and one ***Y*** chromosome, while women have two (2) ***X*** chromosomes [4].
- To experience color blindness, the genetic mutation for color-blindness must be present on the ***X*** chromosome, but for women, this means it must be present on both ***X*** chromosomes.

> Men only need to mutation to be present on their singular ***X*** chromosome, making it much easier for them to inherit color blindness.

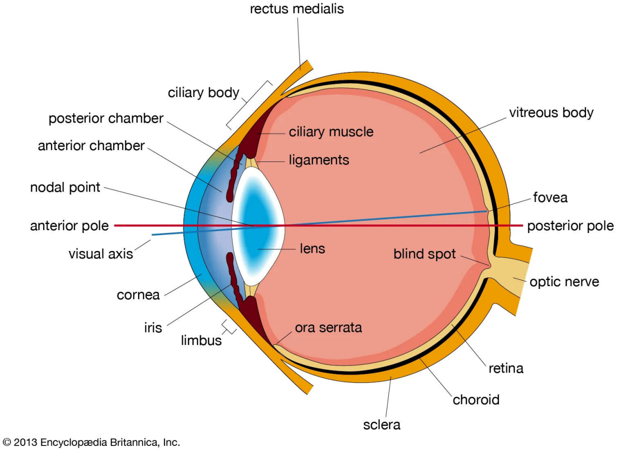


Figure 25: Horizontal section of the eye.

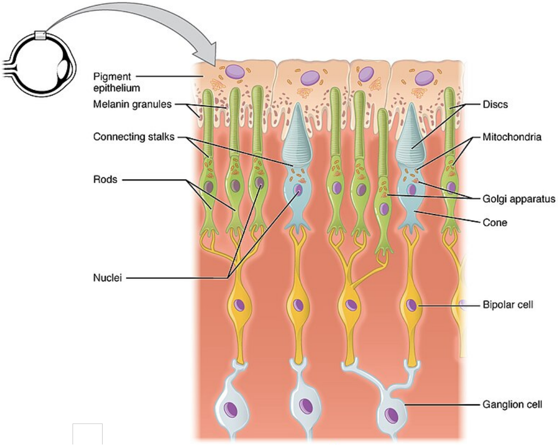


Figure 26: Functional parts of the rods and cones, which are two of the three types of photosensitive cells in the retina.

Trichromacy

- Normal colour vision uses all three (3) types of cone cells.
- Another term for normal colour vision is trichromacy.
- People with normal colour vision are known as trichromats.

Anomalous Trichromacy

- People with faulty trichromatic vision will be colour blind to some extent and are known as anomalous trichromats [3].
- In people with this condition all of their three (3) cone cell types are used to perceive light wavelengths but one type of cone cell perceives light slightly out of alignment.
- There are three (3) different types of effect produced depending upon which cone cell type is faulty and there are also different severities.

- The different anomalous condition types are [5]:

  protanomaly reduced sensitivity to red light,

  deuteranomaly reduced sensitivity to green light (most common),

  tritanomaly reduced sensitivity to blue light (most uncommon).

Achromatopsia

- Can see no colour at all and their world consists of different shades of grey ranging from black to white, rather like seeing the world on an old black and white television set [6].
- Achromatopsia is a specific eye condition in which people see in grey-scale.
- In rare cases, partial Achromatopsia can happen which is a reduced sensitivity to all three (3) cones [6].

Figure 27: Image viewed by someone who has 3 cones.

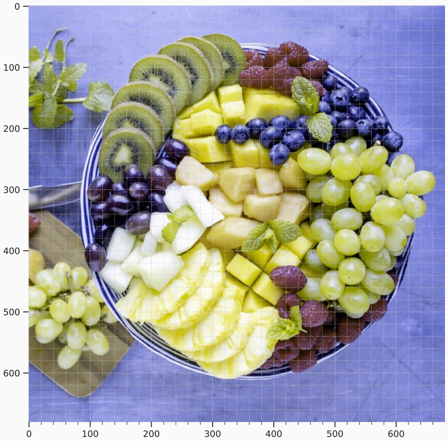


Figure 28: Image viewed by someone who has protanomaly.

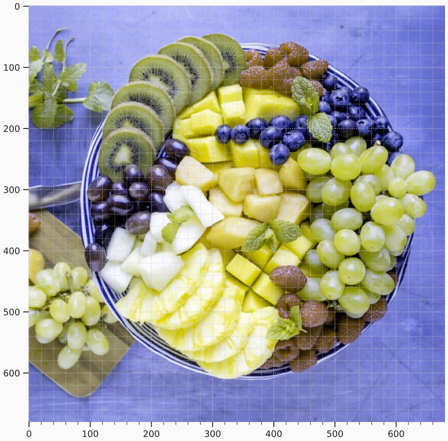


Figure 29: Image viewed by someone who has deuteranomaly.

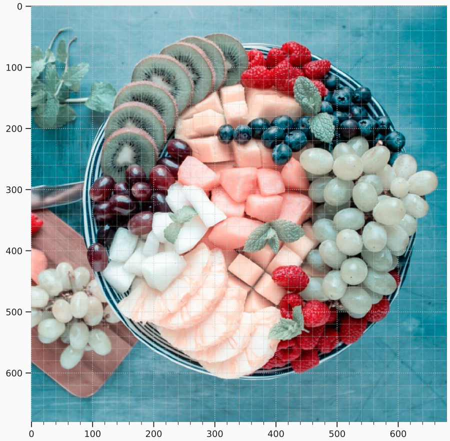


Figure 30: Image viewed by someone who has tritanomaly.

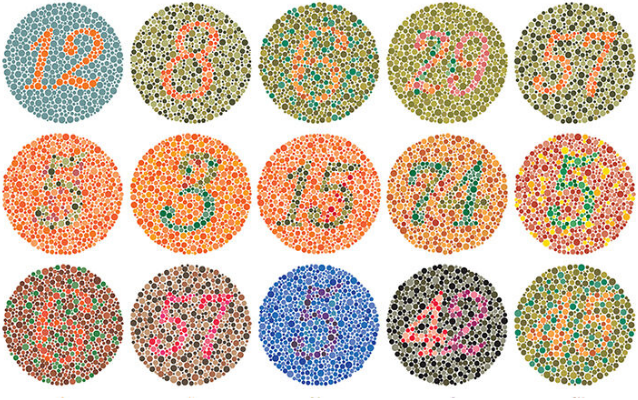


Figure 31: A colour blindness test issued to generally test before taking the driving license [7].

- There are ways to describe the sensitivity of human vision.
- Assume a homogeneous region in an image has an intensity as a function of wavelength (colour) given by $I(\lambda)$.
  - assume $I(\lambda) = I_0$ as a constant.

Wavelength Sensitivity

- The sensitivity of the human eye to light of a certain intensity varies strongly over wavelengths between 380 nm and 800 nm [8].
- Under daylight conditions, human eye is most sensitive at a wavelength of 555 nm, resulting in the fact that green light at produces the impression of highest "brightness" when compared to light at other wavelengths [8].

- The perceived intensity as a function of λ, the spectral sensitivity, for the typical observer is shown below.

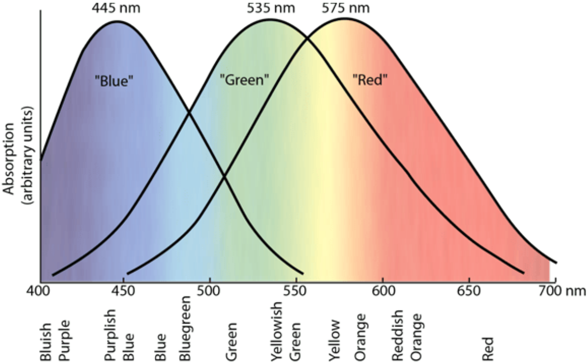


Figure 32: The colour sensitivity of the human eye [9].

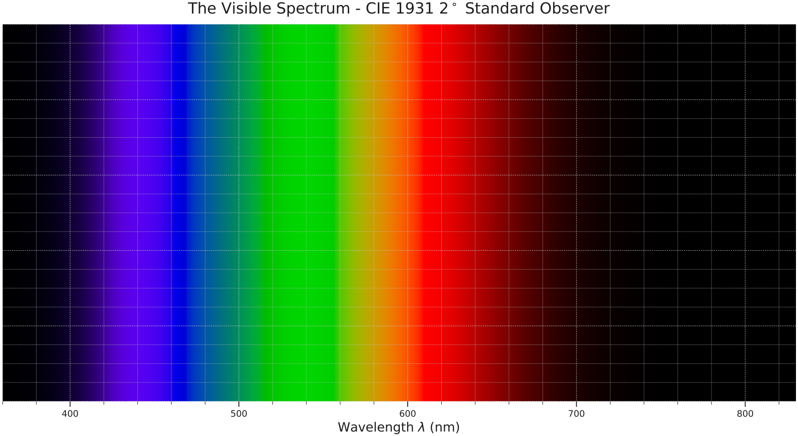


Figure 33: The visible colour spectrum visible with the human eye.

- If the constant intensity (i.e., brightness) I_0 is allowed to vary, then, to a good approximation, the visual response, R, is proportional to the **logarithm** of the intensity.
- This is known as the Weber-Fechner law [10].

> Relates to human perception, specifically the relation between the actual change in a physical stimulus and perceived change.

$$R = \log\left(I_0\right)$$

- This means, equal perceived steps in brightness, $\Delta R = k$, require the physical brightness (i.e., the stimulus) to increase exponentially.

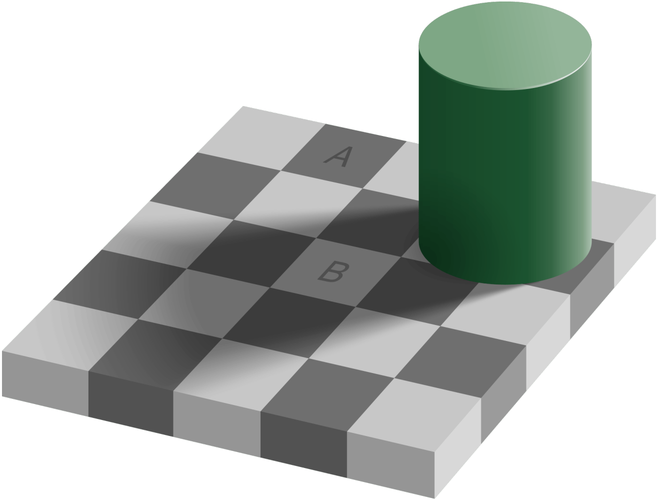


Figure 34: The checker shadow illusion [11].

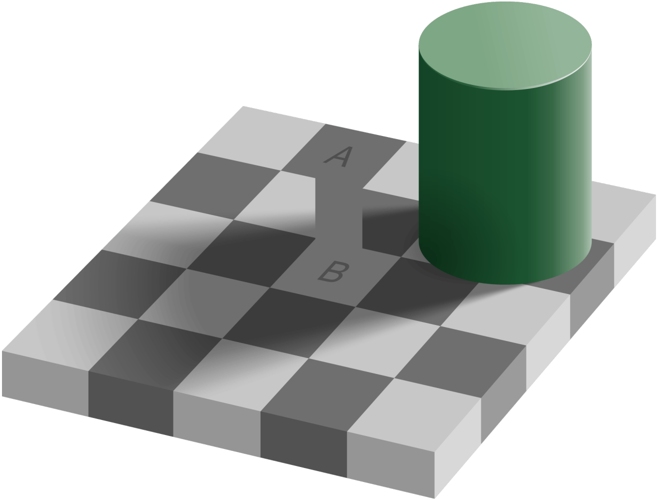


Figure 35: A region of the same shade has been drawn connecting A and B.

Figure 36: A simple picture of a dress divided the internet.

Figure 37: A simple picture of a dress divided the internet [12].

- Human colour perception is complex,
  - We can approximate its behaviour.
- **Standard Observer:** Based on psychophysical measurements, standard curves have been adopted by the CIE as sensitivity curves for the typical observer for three pigments: $\overline{x}(\lambda)$, $\overline{y}(\lambda)$, and $\overline{z}(\lambda)$.

These are not pigment absorption characteristics found in human retina but rather sensitivity curves derived from actual data.

This standard is used by companies to produce monitors and software that are compatible with each other.

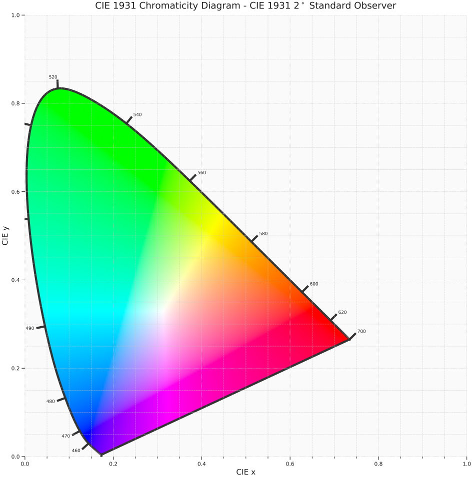


Figure 38: The colour gamut visible to the human eye, standardised by the CIE.

- A standard RGB colour-space defined by both HP and Microsoft in 1996 to use on monitors, printers, and the World Wide Web [13].
- It was subsequently standardised by IEC as IEC 61966-2-1:199 [14].
- sRGB is the **current defined standard colour-space** for the web, and it is usually the assumed colour-space for images that are neither tagged for a colour-space nor have an embedded color profile.
- It codifies the display specifications for the computer monitors in use at the time, which greatly aided its acceptance.
- sRGB uses the same colour primaries and white point as ITU-R BT.709 standard for HDTV, designed to match typical home and office viewing conditions.

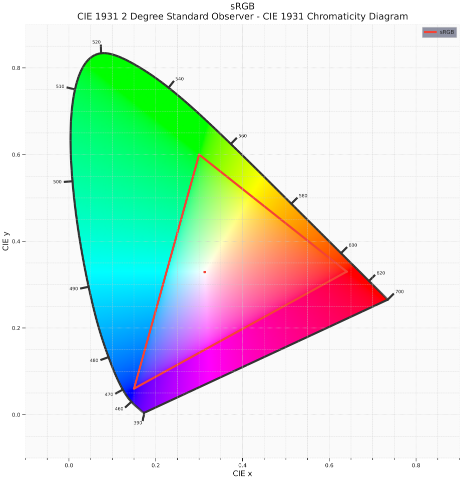


Figure 39: The sRGB colour-space superimposed to the CIE colour-gamut.

- Due to the standardisation of sRGB on the digital-space, and on printers, many low- to medium-end consumer digital cameras and scanners use sRGB as the default working colour-space [15].
- However, consumer-level Charge Coupled Device (CCD)s are typically uncalibrated, meaning that even though the image is being labeled as sRGB, one can not conclude that the image is color-accurate sRGB.

- The wide-gamut RGB colour-space (Adobe Wide Gamut RGB) is developed by Adobe, which offers a large gamut by using pure spectral primary colours [16].
- It is able to store a wider range of colour than sRGB or Adobe RGB.

> For comparison, the wide-gamut RGB colour-space encompasses 77.6% of the visible colours, while Adobe RGB covers 52.1% and sRGB only 35.9% [17].

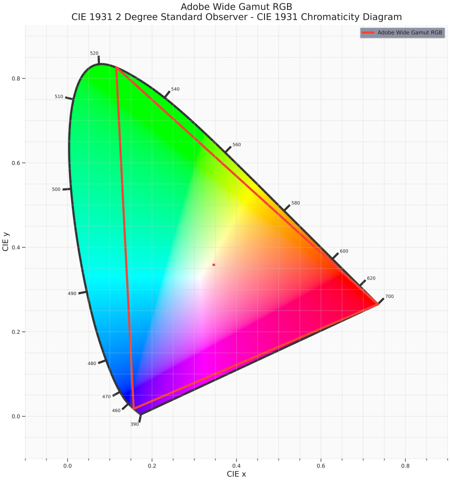


Figure 40: The wide gamut RGB colour-space superimposed to the CIE colour-gamut.

- The ProPhoto RGB colour space, a.k.a. ROMM RGB (Reference Output Medium Metric), is an output referred RGB color space developed by Kodak [18].
- Offers an especially large gamut designed for use with photographic output in mind.
- The gamut encompasses over 90% of possible color space, and 100% of likely occurring real-world surface colours making ProPhoto even larger than the Wide-gamut RGB color space [19].
- The ProPhoto RGB primaries were also chosen in order to minimise hue rotations associated with non-linear tone scale operations.

A downside is that approximately 13% of the visible colours are imaginary colors that do not exist and are impossible colour.

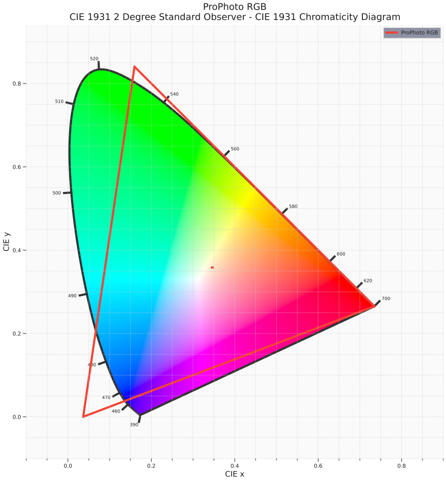


Figure 41: The ProPhoto colour-space superimposed to the CIE colour-gamut.

- The **Adobe RGB (1998)** or **opRGB** is a color space developed by Adobe Inc. in 1998.
- It was designed to encompass most of the colors achievable on CMYK color printers, but by using RGB primary colors on a device such as a computer display.
- The Adobe RGB (1998) color space encompasses roughly 30% of the visible colors specified by the CIE - improving upon the gamut of the sRGB color space, primarily in cyan-green hues.
- It was then standardised by the IEC as IEC 61966-2-5:1999 with a name opRGB (optional RGB color space) and is used in HDMI [20].

- For an arbitrary homogeneous region in an image that has an intensity as a function of wavelength (colour) given by $I(\lambda)$, the three responses are called the tristimulus values:

$$A = \int_{-\infty}^{+\infty} I(\lambda)\ \overline{a}(\lambda)\ d\lambda \quad \text{where} \quad A = \{X,\ Y,\ Z\}\ ,\ a = \{x,\ y,\ z\}\ .$$

- The chromaticity coordinates which describe the perceived colour information are defined as:

$$x = \frac{X}{X+Y+Z}, \quad y = \frac{y}{X+Y+Z}, \quad z = 1 - (x+y)\,.$$

- The tristimulus values are linear in $I(\lambda)$ and thus the absolute intensity information has been lost in the calculation of the chromaticity coordinates $\{x, y\}$.
- All colour distributions, $I(\lambda)$, that appear to an observer as having the same colour will have the same chromaticity coordinates.

- The formulas for converting from the tristimulus values (X, Y, Z) to **RGB** colours (R, G, B) and back are given by:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix} = \begin{bmatrix} 1.19107 & -0.5326 & -0.2883 \\ -0.9843 & 1.9984 & -0.0283 \\ 0.0583 & -0.1185 & 0.8986 \end{bmatrix} \cdot \begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$$

and:

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.6067 & 0.1736 & 0.2001 \\ 0.2988 & 0.5868 & 0.1143 \\ 0.0000 & 0.0661 & 1.1149 \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

- Before we end our look on to these colour spaces, lets have a look at two (2) more standards.

- The **CMYK** model is a subtractive model used in colour printing, and describing the printing process itself.
- The abbreviation CMYK refers to the four inks used:

  cyan, magenta, yellow, and key (black).
- Works by partially or entirely masking colours on a lighter, usually white, background.
- The ink limits the reflected light.
- Such a model is called subtractive because inks subtract the colours red, green and blue from white light.
- White light minus red leaves cyan, white light minus green leaves magenta, and white light minus blue leaves yellow.

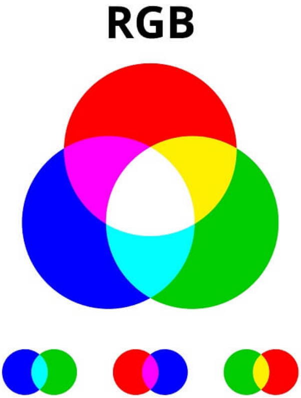


Figure 42: The differences between RGB and CYMK colours [21].

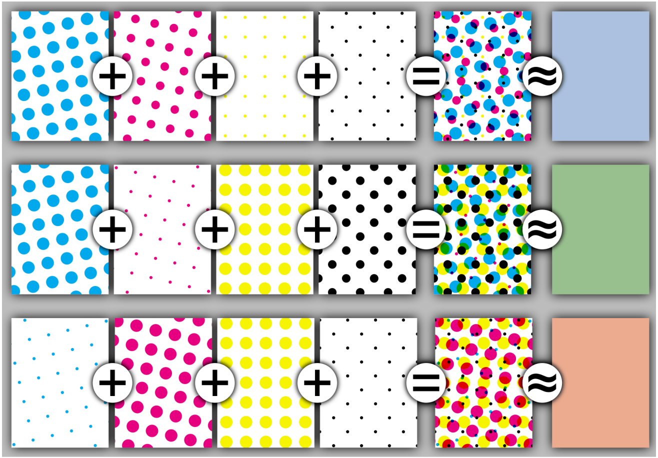


Figure 43: Three examples of color halftoning with CMYK separations, as well as the combined halftone pattern and how the human eye would observe the combined halftone pattern from a sufficient distance [22].

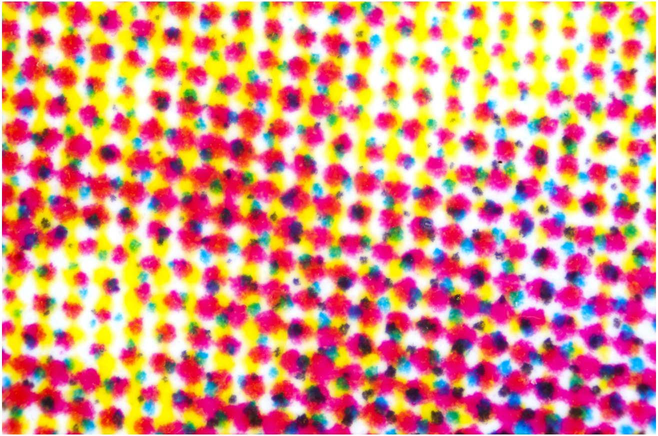


Figure 44: A printer creates any colour by combining dots in particular places relative to the other dots.

- Two most common cylindrical-coordinate representations of points in an RGB color model.
- The two representations rearrange the geometry of RGB in an attempt to be more intuitive and perceptually relevant than the cartesian (cube) representation.
- Developed in the 1970s for computer graphics applications, are used in color pickers, in image editing software, and less commonly in image analysis and computer vision.

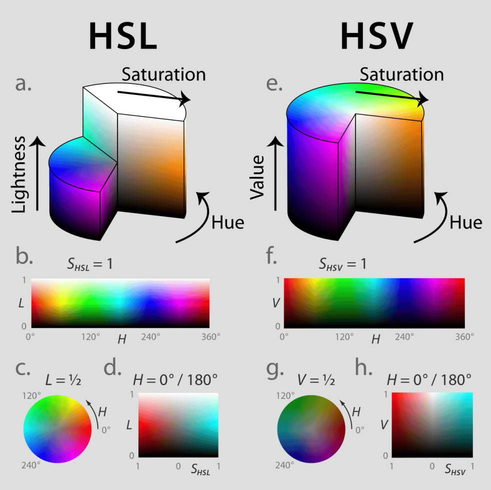


Figure 45: HSL and HSV models.

- $\boldsymbol{YC_bC_r}$ is a family of colour spaces used as a part of the color image pipeline in video and digital photography systems.
- $\boldsymbol{Y}$ is the luma (i.e., brightness) component and CB and CR are the blue-difference and red-difference chroma components.
- Y (with prime) is distinguished from Y, which is luminance, meaning that light intensity is nonlinearly encoded based on gamma corrected RGB primaries.

- CRT uses RGB signals, but they are not the best solution for storing information as they have a lot of redundancy.
- $\boldsymbol{YC_{\mathrm{b}}C_{\mathrm{r}}}$ is a practical approximation, where the primary colours corresponding roughly to red, green and blue are processed into perceptually meaningful information.

- $\boldsymbol{Y'C_bC_r}$ is used to separate out a luma signal (Y) that can be stored with high resolution or transmitted at high bandwidth, and two chroma components (CB and CR) that can be bandwidth-reduced, subsampled, compressed, or otherwise treated separately for improved system efficiency.

> One practical example would be decreasing the bandwidth or resolution allocated to "color" compared to "black and white", since humans are more sensitive to the black-and-white information (see image example to the right). This is called chroma subsampling.

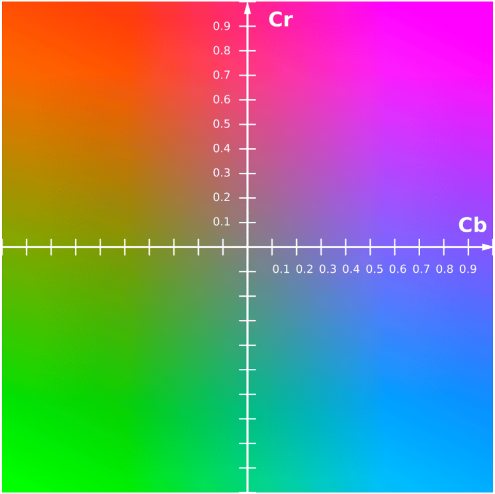


Figure 46: The $Y'C_bC_r$ plane at constant luma Y =0.5 [23].

Image Formats

Table of Contents

(LO1) Types of File Formats used,

(LO2) Defining Compression and its Types,

(LO3) A Look into Selective File Types.

(LO4) A Deep Dive into JPEG Compression.

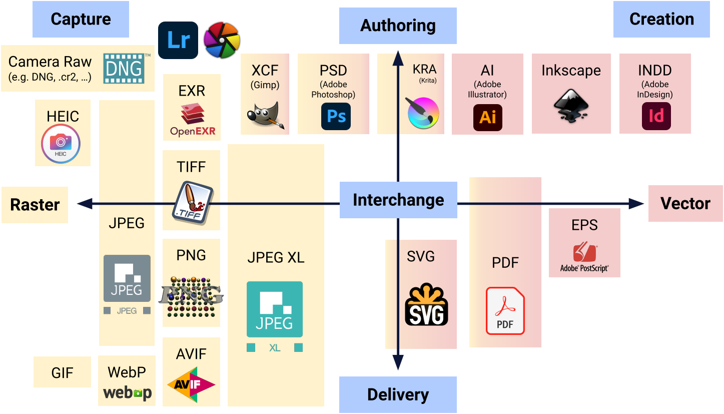


Figure 47: Categorization of image formats by scope which are used in commercial, industrial and personal use [24].

Image Formats

TIFF(`.tif`, `.tiff`) Tagged Image File Format

- Store image data without losing any data.
  - It does not perform any compression on images,
  - High-quality image is obtained however the image size is large,
- Good for printing, and professional work.

JPEG (`.jpg`, `.jpeg`) Joint Photographic Experts Group

- It is a loss-prone (lossy) format.
  - Some data is lost to reduce the image size.
  - Due to compression, some data is lost but that loss is very less
- Used in digital cameras, non-professional prints, e-Mail, etc.

GIF (`.gif`) GIF or Graphics Interchange Format

- Primarily used in web graphics (i.e., Reddit, WhatsApp.)
  - They can be animated and are limited to only 256 colours.
  - These images can also allow transparency.
- GIF files are typically small in size compared to other formats.

PNG (`.png`) PNG or Portable Network Graphics

- It is a lossless image format.
  - It was designed to replace GIF as it supported 256 colours.
  - PNG, whereas, supports 16 million colours.

Bitmap (`.bmp`) Bit Map Image

- It was developed by Microsoft for windows.
- It is same as TIFF due to lossless, no compression property.
- As it is a proprietary format, it is generally recommended to use TIFF.

EPS (`.eps`) Encapsulated PostScript

- It is a vector file.
- EPS files can be opened in applications such as Adobe Illustrator.

Image Formats

RAW Image Files (`.raw`, `.cr2`, `.nef`, `.orf`, `.sr2`)

- Unprocessed and created by a camera or scanner.
- Many DSLR cameras can shoot in RAW.
- These images are the equivalent of a digital negative, meaning that they hold a lot of image information.
- It saves metadata and is used for photography.

- The class of data compression methods using inexact approximations and partial data discarding to represent the content.
- These techniques are used to reduce data size for storing, handling, and transmitting content.
- It can also remove metadata to save up on space [25].
- An example would be JPEG compression.

Data is permanently removed from the file.

Well-designed lossy compression technology often reduces file sizes significantly before degradation is noticed by the end-user.

Image Formats

- A class of data compression that allows the original data to be perfectly reconstructed from the compressed data with no loss of information.
- Lossless compression is possible because most real-world data exhibits statistical redundancy [26].
- In essence, lossless compression algorithms are needed in cases that require compression where we want the reconstruction to be identical to the original.
- Huffman Coding is a perfect example of lossless compression [27].

- A camera RAW image file contains unprocessed or minimally processed data from the image sensor of either a digital camera, a motion picture film scanner, or other image scanner.
- Named RAW as they are not yet processed, and contain large amounts of potentially redundant data.
- Normally, the image is processed by a RAW converter, in a wide-gamut internal color space where precise adjustments can be made before conversion to a viewable format.
- There are dozens of RAW formats in use by different manufacturers of digital image capture equipment.
  - i.e., Nikon uses `.nef` format, and Canon uses `.cr2`.

- It is a lossy compression method.
- Usually stored in the `.jfif` (JPEG File Interchange Format) or the `.exif` (Exchangeable image file format) file format.
- The JPEG filename extension is `.jpg` or `.jpeg`.
- Nearly every camera can save images in the `.jpeg`, which supports eight-bit grayscale images and 24-bit color images
  - Eight bits each for `red`, `green`, `blue`.
- Compression can result in a significant reduction of the file size.
- Applications can determine the degree of compression to apply.
- When not too high, compression does not affect or detract from the image's quality, but JPEG files suffer generation loss when repeatedly edited and saved.

Figure 48: A photo of a European wildcat with the compression rate, and associated losses, decreasing from left to right [28].

Image Formats

- For good lossy compression, throw away the unnecessary information.
- In the case of image compression, Start by understanding which parts of an image are important to human perception, and which aren't.
- Then you find a way to keep the important qualities and trash the rest.
- JPEG, uses two (2) psychovisual principles to compress the image.

1. Changes in brightness are more important than changes in colour:

> Human retina contains about 120 million brightness-sensitive rod cells, but only about 6 million colour-sensitive cone cells.

2. Low-frequency changes are more important than high-frequency changes.

> The human eye is good at judging low-frequency light changes, like the edges of objects. It is less accurate at judging high-frequency light changes, like the fine detail in a busy pattern or texture. Camouflage works in part because higher-frequency patterns disrupt the lower-frequency edges of the thing camouflaged.

Image Formats

- Each pixel is stored as three numbers:
  - Representing red, green and blue.

> The problem, is that the image's brightness information is spread evenly across RGB.

> Remember, brightness is more important than colour.

- We want to isolate brightness from the colour information so we can deal with it separately.
- To do this, JPEG converts the image from RGB to YCbCr.
- In YCbCr all brightness information in one channel (Y) while splitting the colour information between the other two (Cb and Cr).

- Before doing processing, JPEG throws away some of the colour information by scaling down just the Cb and Cr (colour) channels while keeping the important Y (brightness) channel full size.

> This step is optional.

- You can keep all of the colour information, half of it, or a quarter.
- For images, most software will keep half of the colour information;
- for video it is usually a quarter and for this lets do quarter.
- We started with 3 full channels and now we have 1 full channel and 2 x 0.25 channels, for a total of 1.5.
- We are already down to half of the information we started with.

- To use the second observation about perception, start by dividing each of the Y, Cb, and Cr channels up into 8-by-8 blocks of pixels.
- Transform each of these blocks from the spatial domain to the frequency domain.
- Consider just one of these 8-by-8 blocks from the Y channel.
- The spatial domain is what we have now:
  - the value in the upper-left corner represents the brightness (Y-value) of the pixel in the upper-left corner of that block. Likewise, the value in the lower-right corner represents the brightness of the pixel in the lower-right corner of that block. Hence the term spatial: position in the block represents position in the image.

- When we transform this block to the frequency domain, position in the block will instead represent a frequency band in that block of the image. The value in the upper-left corner of the block will represent the lowest-frequency information and the value in the lower-right corner of the block will represent the highest-frequency information.

11	11	10	10	10	10	9	9
11	11	10	10	10	10	9	9
11	11	12	12	9	9	8	8
11	11	12	12	9	9	8	8
10	10	10	10	11	11	9	9
10	10	10	10	11	11	9	9
10	10	8	8	11	11	10	10
10	10	8	8	11	11	10	10

79.5	3.58	-1.39	1.89	0.0	-1.26	0.57	-0.71
0.64	3.91	-0.59	-2.43	0.0	1.62	0.25	-0.78
-0.46	-2.41	2.13	2.09	0.0	-1.4	-0.88	0.48
0.22	-1.61	-0.21	1.34	0.0	-0.9	0.09	0.32
0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
-0.15	1.08	0.14	-0.9	0.0	0.6	-0.06	-0.21
0.19	1.0	-0.88	-0.87	0.0	0.58	0.37	-0.2
-0.13	-0.78	0.12	0.48	0.0	-0.32	-0.05	0.15

Figure 49: On the left we have an input of 8x8 pixels, on the right we have DCT coefficients rounded to two decimal places.

- The next step is to selectively throw away some of the frequency information.
- If you have ever saved a JPEG image and chosen a quality value, this is where that choice comes into play.
- Start with two 8-by-8 tables of whole numbers, called the quantization tables.
- One table is for brightness information, and one is for colour information.
- You will use these numbers on each of the 8-by-8 blocks in the image data by dividing the frequency value in the image data by the corresponding number in its quantization table.

- So the upper-left corner of each 8×8 block in the Y frequency channel will be divided by the number in the upper-left corner of the brightness quantization table, and so on.
- The result of each division is rounded to the nearest whole number and the fractional parts are thrown away.

> The quantization tables are saved along with the image data in the JPEG file. They'll be needed to decode the image correctly.

- If you think carefully about what just happened, you will realize that even though we threw away some frequency information by tossing the decimal parts after division, we still have the same amount of data:
- one number for each pixel from each of the three channels.
- However, this data is now going to be compressed using traditional lossless compression.

- But wait, wasn't the whole reason we used lossy compression in the first place that lossless compression doesn't work well for images? Yes, but that quantization we just did is going to make the data more compressible by making it less noisy. To see why, compare these three number sequences:
- JPEG has one last trick for making the data more compressible: it lists the values for each 8×8 block in a zig-zag pattern that puts the numbers in order from lowest frequency to highest. That means that the most heavily quantized parts (with the largest divisors) are next to each other to make nice, repetitive patterns of small numbers.

Camera

Table of Contents

(LO1) An Overview of Camera Sensors,

(LO2) Camera as a Scientific Instrument,

(LO3) Working with Colour Filters,

(LO4) Lens Technologies.

- The cameras and recording media available for modern digital image processing applications are constantly changing.
- In this section we will focus on two (2) technologies:

  CCD Charge Coupled Device,

  CMOS Complementary Metal-Oxide-Semiconductor,
- The techniques that are used in these technologies are universal.
- There are also variations of this technologies which are more recent:

  EMCCD Electron-multiplying charge-coupled device [29],

  sCMOS Back-illuminated CMOS [30].

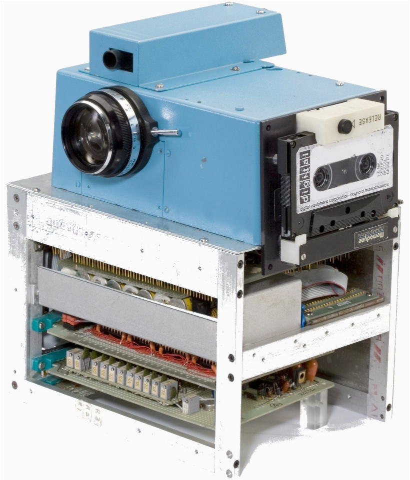


Figure 50: The first digital camera was invented in the 1970s by Kodak [31].

- First sensors used in digital cameras, available since 1970s [32].
- A Charge Coupled Device (CCD) is a highly sensitive photon detector.
- Divided up into a large number of light-sensitive small areas (known as pixels) which can be used to build up an image.
- A photon of light which falls within the area defined by one of the pixels will be converted into one (or more) electrons and the number of electrons collected will be directly proportional to the intensity of the scene at each pixel.

> This allows the camera to be quantitative.

- When the CCD is clocked out, the number of electrons in each pixel are measured and the scene can be reconstructed.

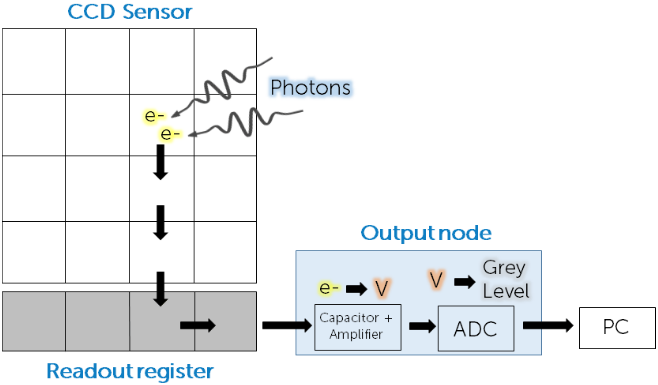


Figure 51: How a CCD sensor works. Photons hit a pixel and are converted to electrons, which are then shuttled down the sensor to the readout register, and then to the output node, where they are converted to a voltage, then grey levels, and then displayed with a PC. [32]

Limitations

- The main issues with CCDs are their lack of speed and sensitivity, making it a challenge to perform low-light imaging or to capture dynamic moving samples.

  Sensor millions of pixels of signal have to be shuttled through one node, creating a bottleneck and slowing the camera.

  Noise to control the read noise from fast moving electrons, CCDs slows down the electrons.

  Refresh Whole sensor needs to be cleared of the electron signal before the next frame can be exposed

- Small full-well capacity limits electron storage in each pixel.
- If a pixel becomes full and displayed the brightest signal, and blooming, where the pixel overflows and the excess signal is smeared down the sensor as the electrons are moved to the readout register.

Figure 52: Examples of blooming caused by saturation of a CCD sensor pixel. Left) Picture of a sunset. The sun is so bright in the image that there is blooming on the sun itself, leaking into the surrounding pixels, and a vertical smear across the whole image. Right) A similar situation with the blooming and smear labeled [32].

- The CMOS image sensor is made up of an array of tiny light-sensitive cells also known as pixels, each of which is connected to a transistor that acts as a switch [33].
- When light strikes the pixels, it is converted into an electrical charge which is amplified and read out by the sensor's readout electronics.
- The output is then converted into a digital signal, ready for further image processing or storage.

> Emerged as an alternative to charge-coupled device (CCD) image sensors and eventually outsold them by the mid-2000s [34].

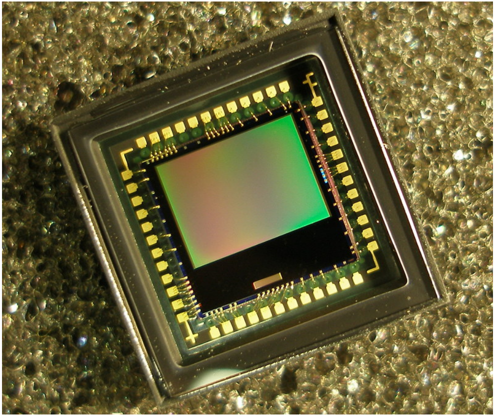


Figure 53: CMOS image sensor [35].

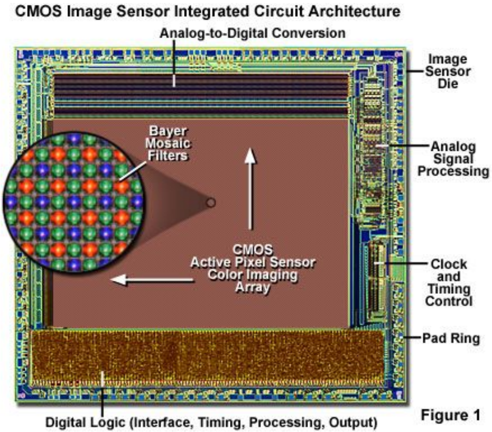


Figure 54: An integrated CMOS Image sensor [36].

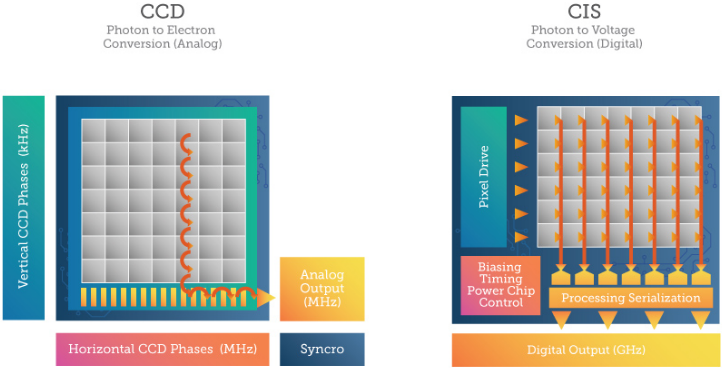


Figure 55: A comparison on the working principle of two types of sensors: CCD and CMOS [37].

The first difference lies in the way the image data is captured/read.

- A CCD captures image data by moving charge packets in a linear sequence from one pixel to the next, which is referred to as charge transfer.
- This method results in improved image quality, but it also has drawbacks such as a slower readout speed (i.e., pixels per second).
- On the other hand, CMOS capture image data by reading out each pixel individually,
  - a process known as parallel readout.

> This results in a faster readout time but generally lower image quality compared to CCD sensors.

- CMOS is cheaper to produce than CCD due to their more complex electronic circuitry and manufacturing process.
- CCD is known to be more power-hungry as compared to CMOS.
  - CCD needs active power as opposed to the CMOS which don't.
- CMOS has a distinct advantage as readout circuits for each pixel is integrated on the same chip as the pixel.
  - This allows for a wide range of enhanced functions such as image processing and signal amplification.

> CCDs are known for their high image quality, lower noise, and greater light sensitivity while CMOSs have the advantage of being faster, less power-hungry and more cost-effective.

- A color filter array (CFA), or color filter mosaic (CFM), is a mosaic of tiny color filters placed over the pixel sensors of an image sensor to capture color information.
- Color filters are needed because the typical photosensors detect light intensity with little or no wavelength specificity and therefore cannot separate color information [38].

> Diazonaphthoquinone (DNQ)-novolac photoresist is one material used as the carrier for making color filters from color dyes or pigments [39, 40].

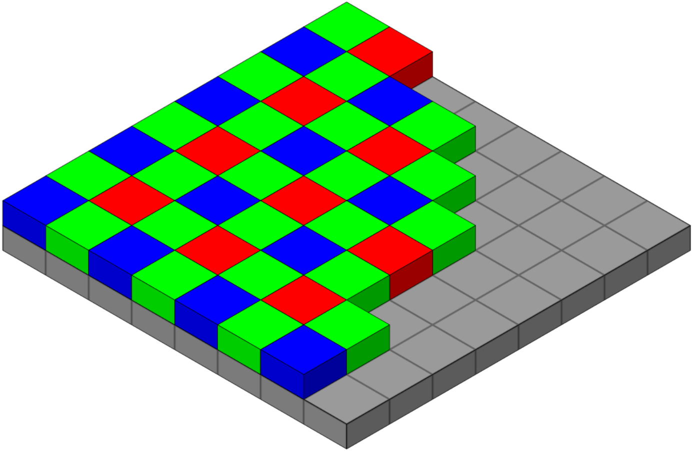
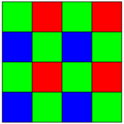
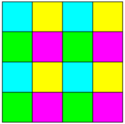


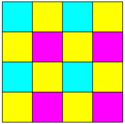
Figure 56: The Bayer color filter mosaic. Each two-by-two submosaic contains 2 green, 1 blue, and 1 red filter, each filter covering one pixel sensor. The reason as to why there are two greens to red and blue is to mimic the physiology of the human eye [41].



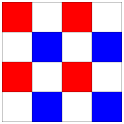
(a) Bayer Filter



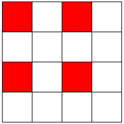
(b) CYGM Filter



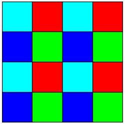
(c) CYYM Filter



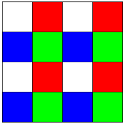
(d) RCCB Filter



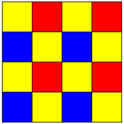
(e) RCCC Filter



(f) RGBE Filter



(g) RGBW Filter



(h) RYYB Filter

Figure 57: Types of colour filter used in commercial and industrial applications

Figure 58: Sony F828, produced in 2003 which uses an RGBE filter. Once the image is taken a demosaicing algorithm is used to reconstruct an image [42].

- It is generally desirable the relationship between the input physical signal and the output signal be linear [43].
- This means that if we have two images, a and b:

$$\mathcal{C} = \mathcal{R}\left\{w_1 a + w_2 b\right\} = w_1 \mathcal{R}\left\{a\right\} + w_2 \mathcal{R}\left\{b\right\}.$$

  where $\mathcal{R}\left(\cdot\right)$ is the camera response and $\mathcal{C}$ is the camera output.
- In practice the relationship between a and $\mathcal{C}$ is given by:

$$\mathcal{C} = \text{Gain} \cdot a^{\gamma} + \text{Offset}.$$

  where γ is the gamma of the recording medium.
- A linear system must have $\gamma = 1$ and $\text{Offset} = 0$.
- Unfortunately, Offset is never zero and must be compensated for if the intention is to extract intensity measurements.

- There are two (2) ways to describe the sensitivity of a camera.

1. We can determine the minimum number of detectable photoelectrons called **absolute sensitivity**.
2. We can describe the number of photoelectrons necessary to change from one digital brightness level to the next called **relative sensitivity**.

Figure 59: Types of camera lenses.

Absolute Sensitivity

- We need a characterisation of the camera in terms of its noise.
- If total noise has a σ of, 100 photoelectrons, then to ensure detectability of a signal we could say that, at the 3σ level (% 99.7), the minimum detectable signal (or absolute sensitivity) would be 300 photoelectrons.
- If all the noise sources, with the exception of photon noise, can be reduced to negligible levels, this means that an absolute sensitivity of less than 10 photoelectrons.
- This can be rephrase to the number of photons needed to have signal equal to noise [44].
- This quantity plays an important role in very-low light applications.

Relative Sensitivity

- When coupled to the linear case, with $\gamma = 1$, leads to the result:

$$\text{Relative Sensitivity} = 1/\text{gain}.$$

The sensitivity or gain can be deduced in two (2) distinct ways.

1. If *a* can be precisely controlled by either shutter time or intensity, the gain can be estimated by estimating the slope of the resulting straight-line curve.
   - To translate this to units, a standard source must be used emitting photons to the camera sensor and the quantum efficiency (η) of the sensor must be known.
2. If the limiting effect of the camera is only the photon noise, then:

$$S = \frac{\mu}{\sigma^2} = \frac{m_c}{s_c}$$

Label		µm × µm	K	e^-/ADU	
C-1	1320 × 1035	6.8 × 6.8	231	7.9	12
C-2	578 × 385	22.0 × 22.0	227	9.7	16
C-3	1320 × 1035	6.8 × 6.8	293	48.1	10
C-4	576 × 384	23.0 × 23.0	238	90.9	12
C-5	756 × 581	11.0 × 5.5	300	109.2	8

Table 4: Sensitivity measurements of sample cameras.

- In a scientific-grade CCD camera (C-1), only 8 photoelectrons (approximately 16 photons) separate two gray levels in the digital representation of the image.
- For a considerably less expensive video camera (C-5), only about 110 photoelectrons (approximately 220 photons) separate two gray levels.

- In a modern camera (CCD or CMOS), the noise is defined by:
  - Amplifier noise in the case of colour cameras.
    - This is more apparent in blue channel.
  - Thermal noise which, itself, is limited by the chip temperature K and the exposure time T,
  - Photon noise which is limited by the photon production rate ρ and the exposure time T.

> We will have a detailed look into these in our lecture called **Noise**.

- Caused by the thermal energy within the camera sensor [45].
- As sensor works, the heat from operation causes thermal energy generation.

These thermal electrons are independent of the photo-electrons.

- Cameras cant decide if the electrons are from the light or thermal.
- Any thermal electrons that accumulated in the sensor pixel wells are counted as signal upon readout, despite not being part of the signal from the sample.
- This is known as the dark current.

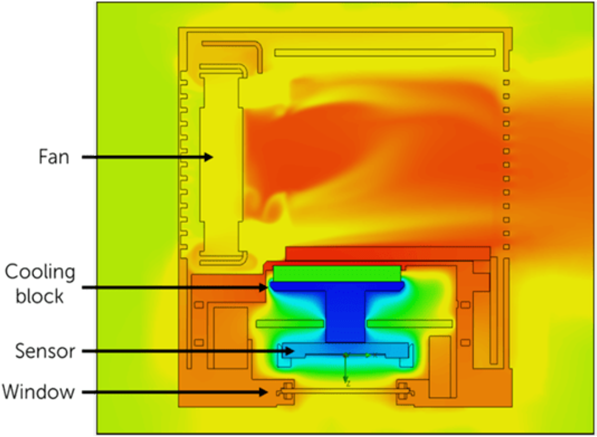


Figure 60: Scientific camera cooling and Citadel Chamber Technology. It shows typical temperature levels within a CMOS camera. The sensor is cooled absolutely and evenly, is an optimal distance from the window, and heat is effectively expelled (edited from [45]).

Prime 95B sCMOS (Normal Operation)	0.55
Prime 95B sCMOS (Air-Cooled)	0.3
Retiga R6 CCD	0.00073
Retiga E7 CMOS	0.001
Nikon D5300 (ISO 200)	6.23

Table 5: Values of Dark Current on Cameras.

- Quantum efficiency (QE) is the measure of the effectiveness to convert incident photons into electrons.

> For example, if a sensor had a QE of 100% and was exposed to 100 photons, it would produce 100 electrons of signal.

- In practice, sensors are never 100% efficient, and different sensor technologies have different QE values.

> The highest-end scientific cameras can achieve up to 95% QE but this is dependent on the wavelength of light being detected.

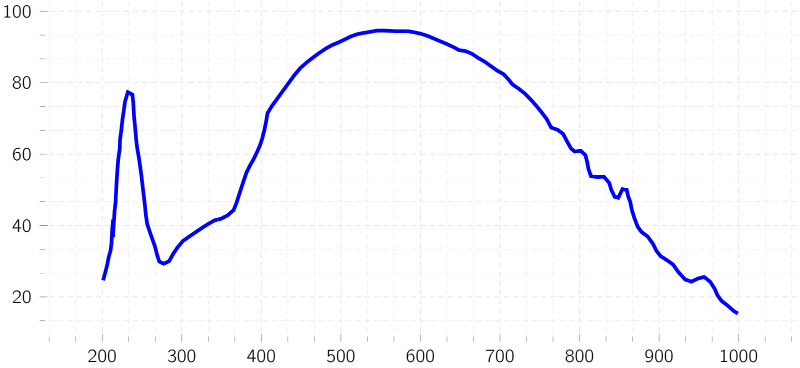


Figure 61: The quantum efficiency (QE) of a 95% quantum efficient sensor at different photon wavelengths. 95% QE is possible at 500-600 nm wavelengths (green/yellow) but it is less efficient at shorter (violet, 300-400 nm) and longer (infrared, 800-1000 nm) wavelengths. This particular sensor also has a peak in QE at near-UV wavelengths, around 220-250 nm.

- ISO is a camera setting which brightens or darkens a photo.
- Increasing the ISO number, makes photos progressively brighter.
- For that reason, ISO can help you capture images in darker environments, or be more flexible about your aperture and shutter speed settings.

> Raising your ISO has consequences. A photo taken at too high of an ISO will show a lot of grain, also known as noise, and might not be usable.

- Brightening a photo via ISO is always a trade-off.
- Only raise your ISO when you are unable to brighten the photo via shutter speed or aperture instead.

Figure 62: A comparison of different ISO numbers. The ISO controls how sensitive the camera is to the light and therefore a higher ISO number would indicate a brighter image [46].

Camera

- A typical lens used with an industrial camera is actually a lens system made of multiple types of optical lenses within an enclosure.
- This type of camera lens is technically called a "compound lens."
- However, a camera lens is the colloquial term for a lens system when it is built with a mounting ring that can fit a variety of cameras.
- Choosing the correct camera lens for a vision system is critical for achieving a specific imaging result.

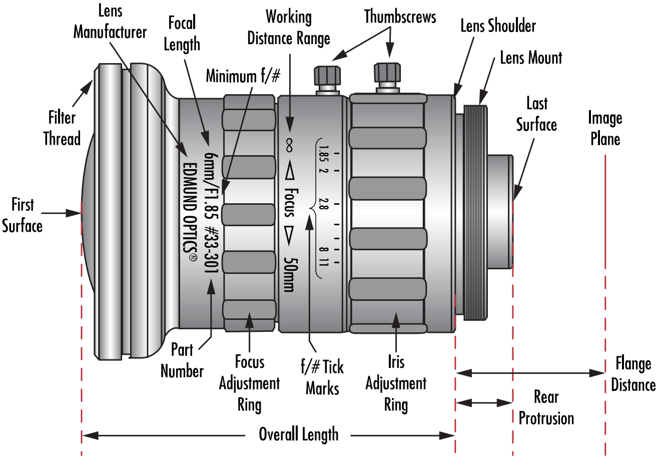


Figure 63: Anatomy of a lens [47].

Figure 64: Different apertures of a lens.

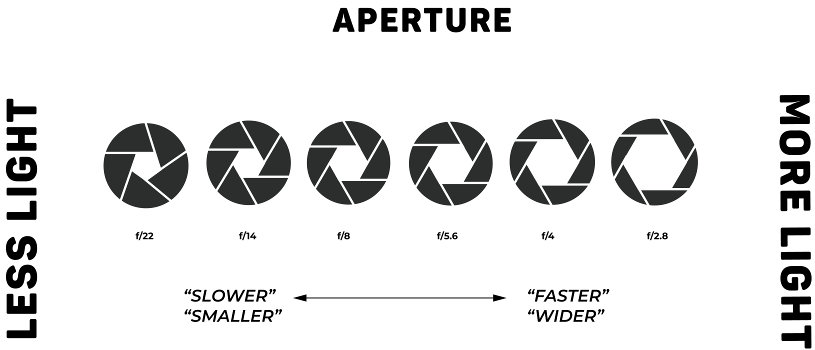


Figure 65: The effects of aperture on the image.

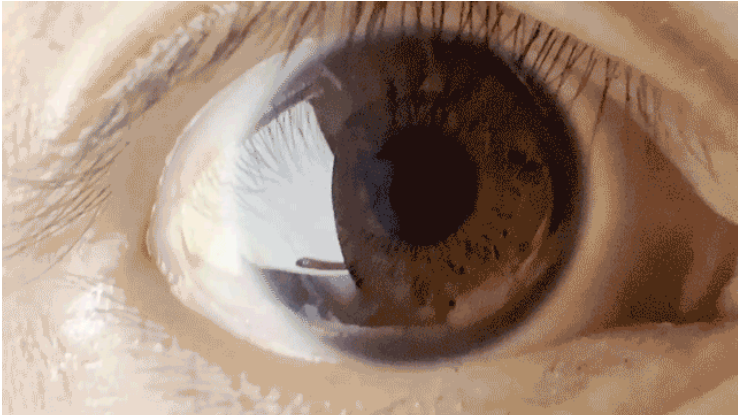


Figure 66: The entrance pupil is typically about 4 mm in diameter, although it can range from as narrow as 2 mm (f/8.3) in diameter in a brightly lit place to 8 mm (f/2.1)

- Choosing the proper lens needs careful consideration focal length.
- Focal length is the distance between the camera's sensor and where the light in the camera lens is focused.
- It is often measured in millimetres (mm) and provides plenty of information about what type of image will be captured,
  - such as the angle of view (often referred to as the field of view or FoV)
  - the magnification of the target within the image.

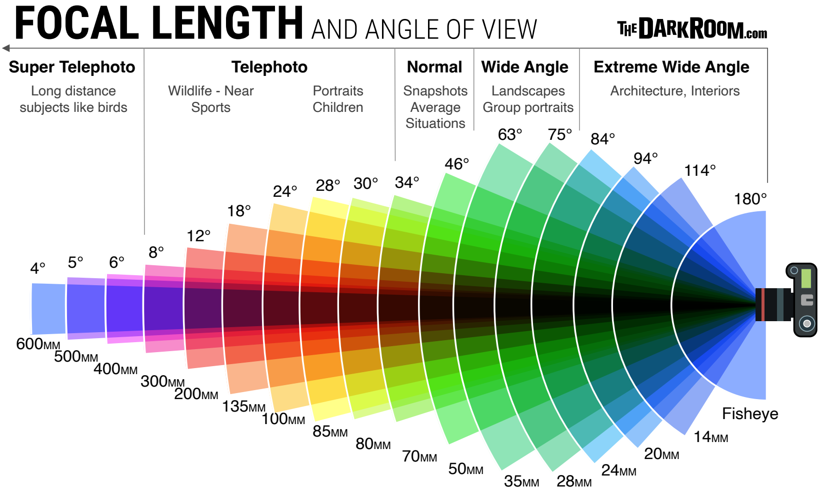


Figure 67: The effects of focal length and the angle of view [48].

- There are two (2) main types of camera lenses:
    1. Prime lens,
    2. Zoom lens (also called varifocal/kit)

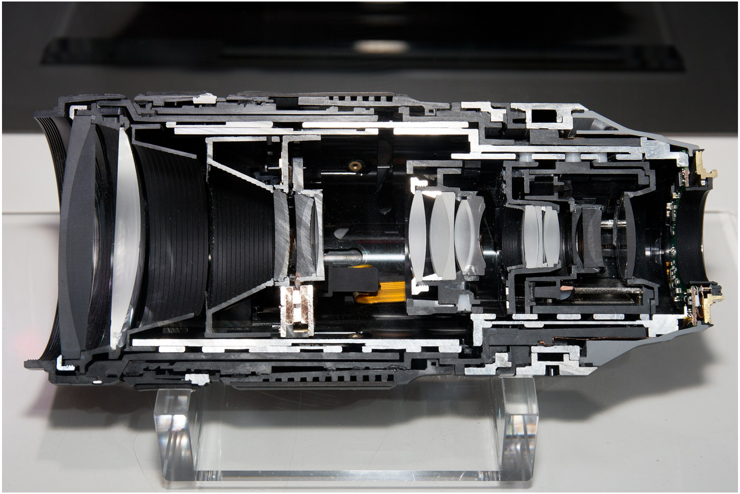


Figure 68: The cross-section of Fujinon XF100-400mm zoom lens [49].

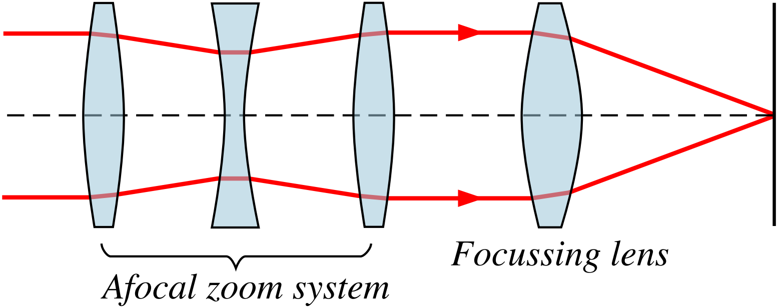


Figure 69: A simple zoom lens system. The three lenses of the afocal system are L_1, L_2, L_3 (from left). L_1 and L_2 can move to the left and right, changing the overall focal length of the system [50]

- An important issue in zoom lens design is the correction of optical aberrations (such as chromatic aberration and, in particular, field curvature) across the whole operating range of the lens;
- This is considerably harder in a zoom lens than a fixed lens, which needs only to correct the aberrations for one focal length.

Figure 70: Photographic example showing a high quality lens (top) compared to a lower quality one exhibiting transverse chromatic aberration [51].

- Prime lens is a fixed focal length photographic lens (as opposed to a zoom lens), typically with a maximum aperture from f2.8 to f1.2.
- While a prime lens of a given focal length is less versatile than a zoom lens, it is often of superior optical quality, wider maximum aperture, lighter weight, and smaller size.
- These advantages stem from having fewer moving parts, optical elements optimized for one particular focal length, and a less complicated lens systems that creates fewer optical aberration issues.
- Larger maximum aperture (smaller f-number) facilitates photography in lower light, and a shallower depth of field.

Display

Table of Contents

Learning Outcomes

(LO1) Refresh Rate,
(LO2) Display Technologies,
(LO3) TV Standards,
(LO4) Resolution.

- Defined as the number of complete images that are written to the screen per second.
- For video, TV refresh rate is either 25 (PAL)[52] or 29.97 (NTSC) images.
- For computer displays, the refresh rate can vary with common values being 30 Hz, 60 Hz and 144 Hz.
- At values above 60 Hz visual flicker is negligible at virtually all illumination levels.
- To prevent the appearance of visual flicker at refresh rates below 60 Hz, the display can be interlaced.
- Standard interlace for video systems is 2:1.
- Since interlacing is not necessary at refresh rates above 60 Hz, an interlace of 1:1 is used with such systems.

Display

It's important to remember images are seen in the first place.

1. Light passes through the cornea at the front of your eye until it hits the lens.
2. The lens then focuses the light on a point at the very back of the eye in a place called the retina.
3. Then, photo-receptor cells at the back of your eye turn the light into electrical signals, while the cells known as rods and cones pick up on motion.
4. The optic nerve carries the electrical signals to your brain, which converts the signals into images.

What happens when watching something with high FPS rate.

- Are you actually seeing all those frames that flash by?
  - After all, your eye doesn't move as fast as 30 motions per second [53].

The short answer is that you may not be able to consciously register those frames, but your eyes and brain may be aware of them.

- 60 FPS rate is generally accepted as the uppermost limit.
- Some research suggests that your brain might actually be able to identify images that you see for a much shorter period of time than experts thought.
- For example, in [54], it was observed the brain can process an image that your eye sees for only 13 milliseconds.
- That's especially rapid when compared with the accepted 100 milliseconds that appears in earlier studies.

Higher FPS also has less stress (i.e., less blinking) on the eye [55].

Thirteen milliseconds translate into about 75 frames per second.

Display

- Interlacing is a method used to create images on a display, like a TV or computer screen [56].
- In an interlaced display, the picture is made by scanning alternating lines [57].
- It scans every other line first, and then it fills in the missing lines in the next scan.
- This method lets the screen refresh faster and at a lower cost.
- A big downside is the picture might flicker or show visible lines.

Figure 71: An Example of Interleaving during a live-stream event [58].

Faster Refresh Rate

- Interlacing allows higher refresh rate as it only needs to update half of the lines on the screen at any one time.
  - Images update quickly, making videos appear smoother.

Reduced Bandwidth

- Interlacing transmits only half the image data at a time.
- Advantageous when streaming video content over slow internet connections as it can start displaying content more quickly,
  - even though the full quality might be achieved a bit later.

Cost Efficiency

- Displays using interlacing are often less expensive to produce.

- A common problem is flickering, due to updating alternate lines instead of the whole image at once.
- This can make the picture seem unstable or shaky, noticeable when there are fast-moving scenes or during scrolling text.
- Another issue is the appearance of visible lines or combing.
- Occurs when the alternating lines don't blend perfectly, making it look like there are gaps or lines through moving objects.
- Artefacts can be distracting and reduces clarity and quality.
- As only half of the image is displayed at a time, interlaced videos can appear less sharp and detailed compared to progressive scan videos, where each frame shows the full image.

> Makes Interlaced content seem outdated or lower quality, especially on modern high-resolution displays where every detail counts.

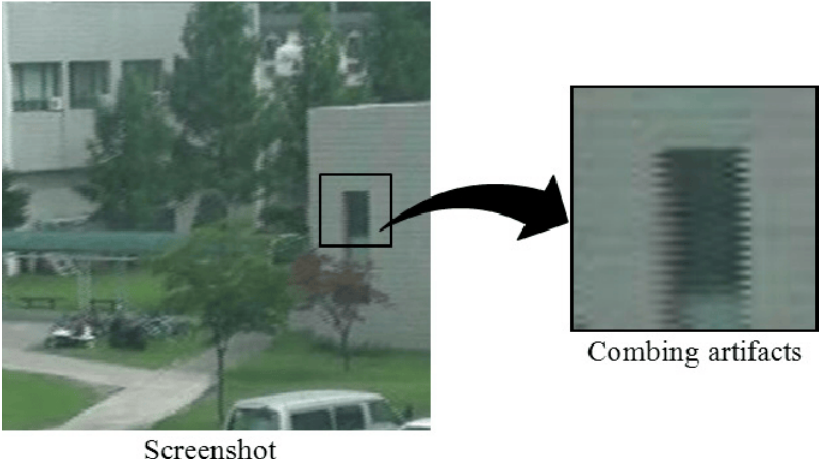


Figure 72: An example of the combing effect observed in an interlaced video [57].

Figure 73: A monochrome CRT as seen inside a Macintosh Plus computer [59].

- An LED emits light as a result of electric luminescence and is one of the most energy-efficient and power-saving ways to produce light.
- Consists of solid materials without movable parts and is often moulded into transparent plastic.
- An LED display consists of many closely-spaced LEDs.
  - By varying the brightness of each LED, the diodes jointly form an image.
- To create a bright colour image, additive colour mixing are used.
- By adjusting the intensity of the diodes, billions of colours can be formed.

> When you look at the LED screen from a certain distance, the array of coloured pixels are seen as an image.

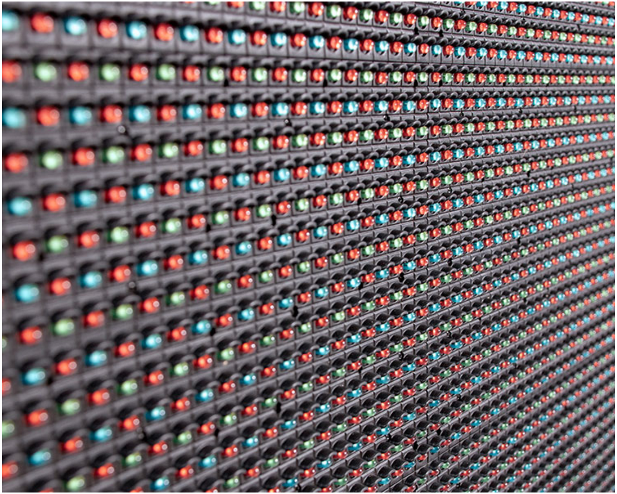


Figure 74: From a close-up it can be seen a display is made from RGB diodes [60].

Display

- LCDs are lit by a backlight, and pixels are switched on and off electronically while using liquid crystals to rotate polarized light.
- A polarising glass filter is placed in front and behind all the pixels, the front filter is placed at 90 degrees.
- In between both filters are the liquid crystals, which can be electronically switched on and off.
- By controlling the electric source the orientation of the liquid crystals can be finely controlled which in turn allows specific light to pass through the second polarising filter.

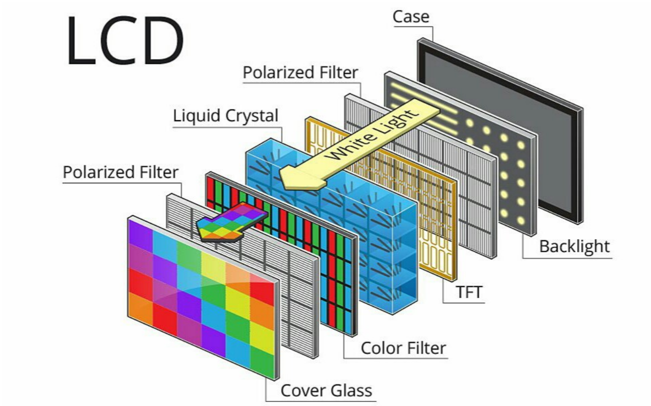


Figure 75: An exploded view of the LCD technology [61].

NTSC (National Television Standard Committee)

- Analog colour-encoding video system used in DVD players and, until recently, television broadcasting in North America.
- In 1953 a new TV standard was introduced by the National Television System Committee and named “NTSC.”
- This format was developed with the intention to be compatible with most TV sets in the country, whether color or black-and-white.
- Even though modern television broadcasters switched to digital, the number of resolution lines and the frame rate they use are the same as established by the NTSC format.

PAL (Phase Alternating Line)

- Another video mode system for analog color television, also used in DVD and Blu-ray players.
- Designed in the late 1950s in Germany, the PAL format was supposed to deal with certain weaknesses of NTSC, including signal instability under poor weather conditions, which was especially relevant for European broadcasters.
- The new standard was to solve the problem by reversing every other line in a TV signal and thus eliminating errors.
- PAL also provided the locally required picture frequency - 50 Hz.
- This format, unlike NTSC, is still employed for broadcasting in the countries where it was adopted.

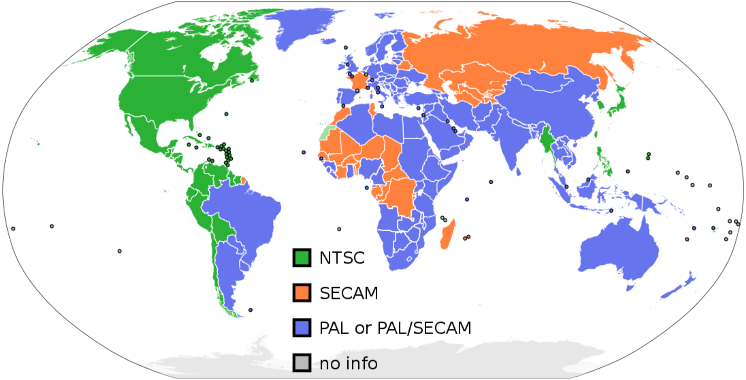


Figure 76: Analog television encoding systems by nation; NTSC (**green**), SECAM (Séquentiel de couleur à mémoire) (**orange**), and PAL (**blue**)

Figure 77: The Philips circle pattern was a physical card which a television camera was pointed, allowing for simple adjustments of picture quality. Such cards are still often used for calibration, alignment, and matching of cameras and camcorders [62].

Display

- The pixels stored in computer memory, although they are derived from regions of finite area in the original scene, may be thought of as mathematical points having no physical extent.
- When displayed, the space between the points must be filled in.
- The brightness profile of a CRT spot is approximately Gaussian and the number of spots that can be resolved on the display depends on the quality of the system.
- It is relatively straightforward to obtain display systems with a resolution of 72 spots per inch (28.3 spots per cm.)
- This number corresponds to standard printing conventions.

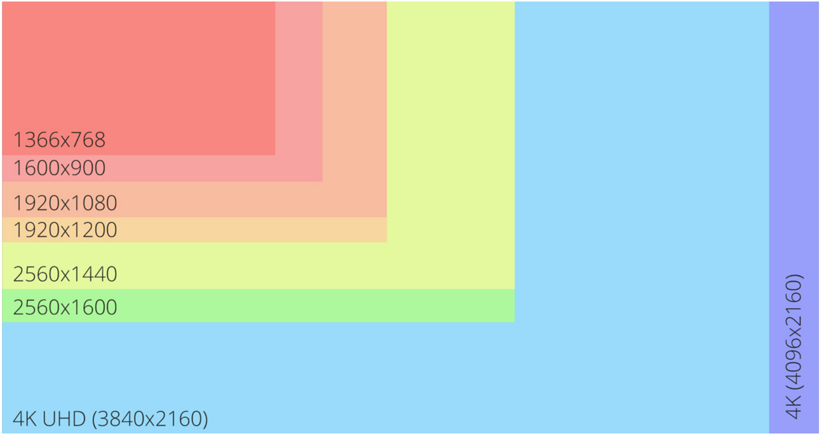


Figure 78: A comparison of different resolution standards [63].

Noise

Table of Contents

- Images acquired through modern sensors may be contaminated by a variety of noise sources.

> Noise is of stochastic variations, opposed to being deterministic.

- We assume to be dealing with modern sensors (CCD, CMOS) where photons produce electrons or referred as photo-electrons.
- Nevertheless, most observations we make about noise and its various sources hold equally well for other imaging modalities.

> Modern technology has reduced the noise levels associated with various electro-optical devices to almost negligible levels except one which **forms the absolute limiting case**.

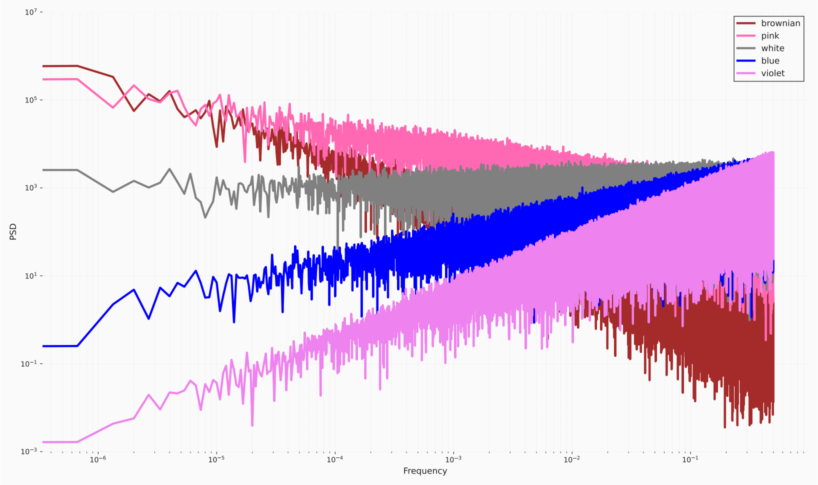


Figure 79: A comparison of different colours of Noise and their power vs. frequency.

Noise

- White noise is a signal, named by analogy to white light, with a flat frequency spectrum.
- For example, with a white noise signal, the range of frequencies between 40 Hz and 60 Hz contains the same amount of power as the range between 400 Hz and 420 Hz,
    - As both these intervals are 20 Hz wide.

Noise

Let's generate some white noise. First is to define `noise_psd`

C.R. 7

python

```python
def noise_psd(N, psd = lambda f: 1):
        X_white = np.fft.rfft(np.random.randn(N));
        S = psd(np.fft.rfftfreq(N))
        # Normalize S
        S = S / np.sqrt(np.mean(S**2))
        X_shaped = X_white * S;
        return np.fft.irfft(X_shaped);
```

Continuing on define a template function which will be inherited.

C.R. 8

python

```python
def PSDGenerator(f):
    return lambda N: noise_psd(N, f)
```

The aforementioned code uses anonymous functions.

C.R. 9

python

```python
@PSDGenerator
def white_noise(f):
    return 1;
```

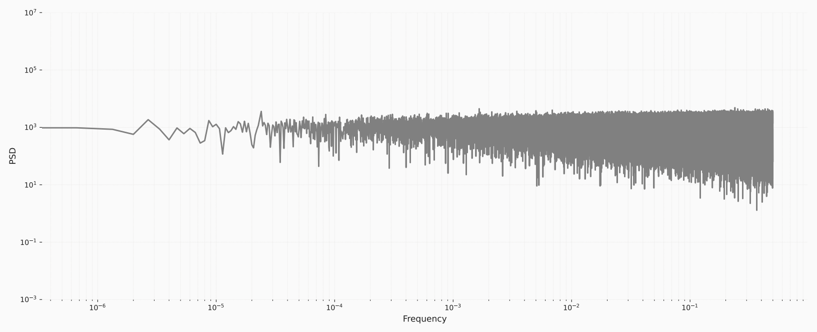


Figure 80: PSD of white noise.

- Pink noise's PSD decreases 3.01 dB per octave with increasing frequency (i.e., $\propto 1/f$).
- Each octave interval (halving or doubling in frequency) carries an equal amount of noise energy.

> In audio form, Pink noise sounds like a waterfall.

- It is used in tuning loudspeaker systems in professional audio [64].
- It is often observed signals in biological systems.
    - i.e., fluctuations in tide and river heights, quasar light emissions, heart beat.

C.R. 10

python

```python
@PSDGenerator
def pink_noise(f):
    return 1/np.where(f == 0, float('inf'), np.sqrt(f))
```

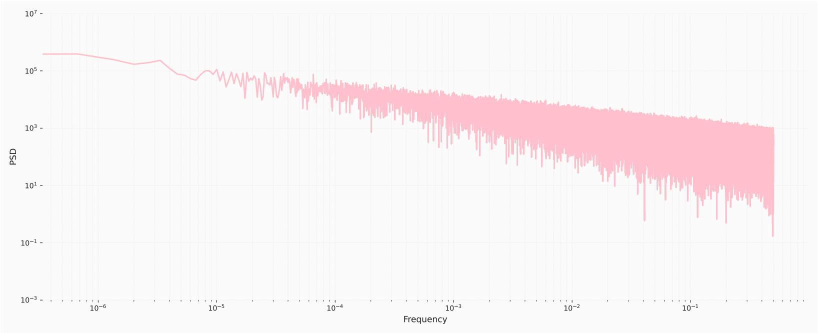


Figure 81: PSD of pink noise.

- Brown noise[1], has a PSD which decreases 6.02 dB per octave with increasing frequency (i.e., $\propto 1/f_2$).
- i.e., a running washing machine, fan noise from an air conditioner or ventilation system.
- Brown noise can be produced by **integrating white noise**.

Brownian noise can also be computer-generated by first generating a white noise, applying Fourier-transforming, then dividing the amplitudes of the different frequency components by the frequency, or by the frequency squared.

C.R. 11

python

```python
@PSDGenerator
def brownian_noise(f):
    return 1/np.where(f == 0, float('inf'), f)
```

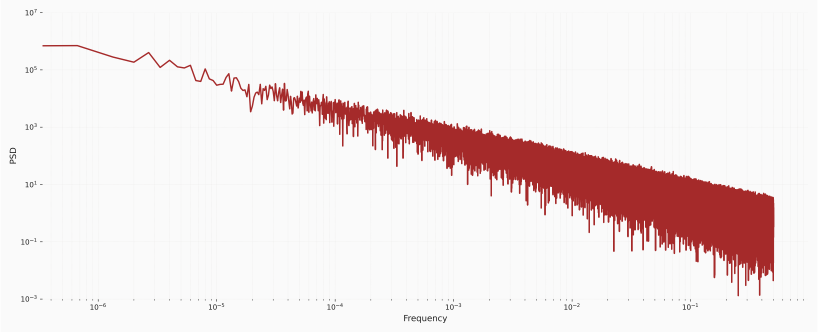


Figure 82: PSD of brown noise.

- Also called purple noise.
- It has a PSD which increases 6.02 dB per octave with increasing frequency (i.e., $\propto f^2$).
- GPS acceleration errors are an example of violet noise processes as they are dominated by high-frequency noise.
- It is also known as differentiated white noise, due to its being the result of the differentiation of a white noise signal.

C.R. 12

python

```python
@PSDGenerator
def violet_noise(f):
    return f;
```

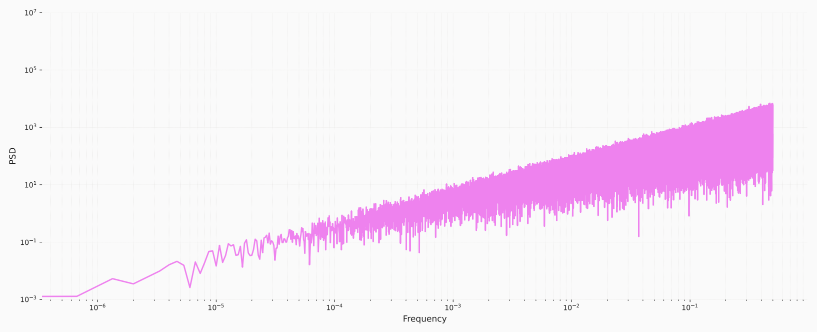


Figure 83: PSD of violet noise.

- Also called azure noise.
- It has a PSD which increases 3.01 dB per octave with increasing frequency (i.e., $\propto f$).

> In computer graphics, the term blue noise is sometimes used more loosely as any noise with minimal low frequency components and no concentrated spikes in energy which is used in for dithering.

C.R. 13

python

```python
@PSDGenerator
def blue_noise(f):
    return np.sqrt(f);
```

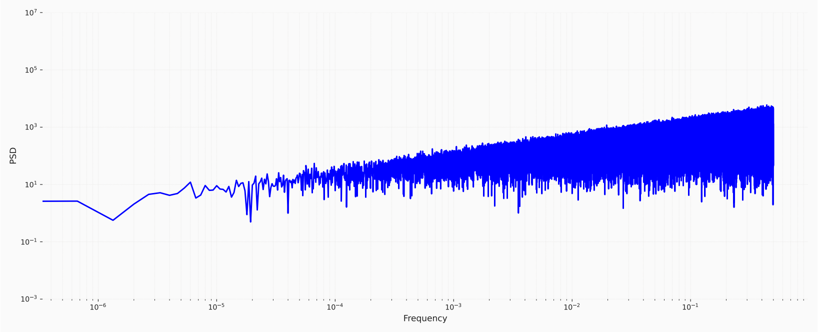


Figure 84: PSD of blue noise.

Dithering

- An intentional noise to randomise quantization error.
- Used in processing of both digital audio and video data, and is often one of the last stages of mastering audio to a CD.
- A common use is converting a grey-scale image to black and white, so black dot density approximates the average grey level.

Figure 85: Image on left is original. Center image reduced to 16 colours. Right image also 16 colors, but dithered to reduce banding effect.

Figure 86: An example of colour banding, visible in the sky.

- When the signal is based upon light, then the quantum nature of light plays a significant role.
- A single photon at $\lambda = 500\,\text{nm}$ carries an energy of:

$$E = h\upsilon = hc/\lambda = 3{,}97 \times 10^{-19}\,\text{J}.$$

- Nowadays, CCDs cameras are able to count individual photons.

The problem comes from the statistical nature of photon.

We cannot assume that, in a given pixel for two consecutive but independent observation intervals of length T, the same number of photons will be counted.

- Photon production is governed by quantum physics, restricting us to talking about an average number of photons within a window.

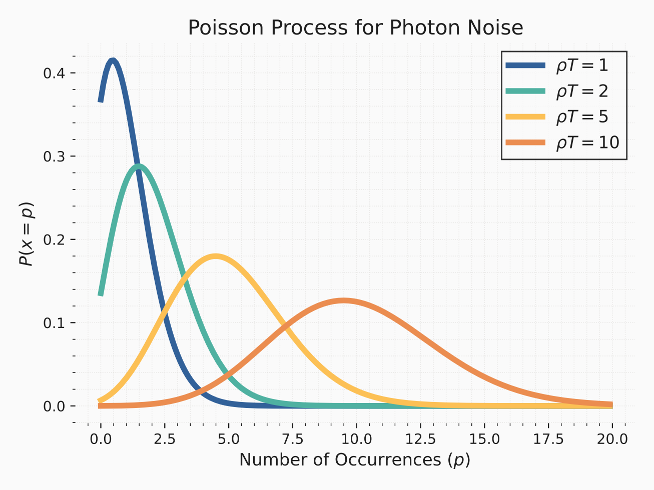


Figure 87: The distribution of the photon noise which is a Poisson process.

- The Poisson distribution is named after Simeon-Denis Poisson.
- Unlike a lot of distributions, it only has one (1) parameter: θ.

This parameter must be positive.

- The formula for the probability density function (PDF) is:

$$P\left(X = x\right) = \frac{e^{-\theta}\theta^{x}}{x!}, \quad \text{for} \quad x = 0, 1, 2, 3, \ldots$$

Example

Consider a computer system with Poisson job-arrival stream at an average of 2 per minute. Determine the probability that in any one-minute interval there will be:

(i) 0 jobs,

(ii) exactly 2 jobs,

(iii) at most 3 arrivals,

(iv) What is the maximum jobs that should arrive one minute with 90 % certainty

Solution

(i) No Job Arrivals

$$P(X=0)=e^{-2}\approx .135 \quad \blacksquare$$

(i) Exactly 3 Job Arrivals

$$P(X=3)=e^{-2}\frac{2^3}{3!}\approx .18 \quad \blacksquare$$

Solution

(iii) At most 3 Arrivals

$$
\begin{aligned}
P\left(X \leq 3\right) &= P\left(0\right) + P\left(1\right) + P\left(2\right) + P\left(3\right), \\
&= e^{-2} + e^{-2}\frac{2}{1} + e^{-2}\frac{2^2}{2!} + e^{-2}\frac{2^3}{3!} \\
&= .135 + .270 + .270 + .180 \\
&= 0.857 \quad \blacksquare
\end{aligned}
$$

For more than 3 arrivals:

$$
\begin{aligned}
P\left(X > 3\right) &= 1 - P\left(X \leq 3\right) \\
&= 1 - 0.857 \\
&= 0.142 \quad \blacksquare
\end{aligned}
$$

Noise

- The probability distribution for p photons in an observation window of length T seconds is known to be Poisson:

$$P(p|\rho,\ T) = \frac{(\rho T)^p \, e^{-\rho T}}{p!}$$

  where ρ is the rate of photons per second.

- Even if there were no other noise sources in the imaging chain, the statistical fluctuations of photon counting over a finite time interval T would still lead to a finite signal-to-noise ratio (SNR).
- Rewriting the SNR, average (μ) value and standard deviation (σ) are given by:

$$\mu = \rho T \quad \sigma = \sqrt{\rho T} \qquad\qquad \mathrm{SNR} = 10 \log_{10}(\rho T) \quad \mathrm{dB}$$

Noise

- Traditional assumptions about signal and noise do not hold:
  - photon noise is not independent of the signal,
  - photon noise is not Gaussian,
  - photon noise is not additive.
- For bright signals, where ρT exceeds 10^5, the noise fluctuations due to photon statistics can be ignored if the sensor has a sufficiently high saturation level.

> If a pixel can be thought of as a well of electrons, saturation refers to the condition where the well becomes filled. The amount of charge that can be accumulated in a single pixel is determined largely by its area. However, due to the nature of the potential well, which holds charge within a pixel, there is less probability of trapping an electron within a well that is approaching saturation.

- Electrons can be freed from the CCD through thermal vibration.
- These freed electron are indistinguishable from true photoelectrons.
- Can be caused by quantum effects (i.e., quantum tunnelling)
- By cooling the CCD chip it is possible to significantly reduce the number of thermal electrons causing thermal noise or dark current.

Noise increases exponentially with temperature.

- A relatively small electric current flowing through photosensitive devices even when no photons enter the device.
- It consists of the charges generated in the detector when no outside radiation is entering the detector.
- It is referred to as reverse bias leakage current in non-optical devices and is present in all diodes.
- Physically, dark current is due to the random generation of electrons and holes within the depletion region of the device.

- Increasing the exposure[1] increases the number of thermal electrons.
- The probability distribution of thermal electrons is also a Poisson process where the rate parameter is an increasing function of temperature.
- There are techniques for suppressing dark current such as estimating the average dark current.
    - i.e., for the given exposure and then subtracting this value from the CCD pixel values before the A/D converter.
- While this does reduce the dark current average (i.e., μ), it does not reduce the dark current standard deviation (σ).
    - This also reduces the possible dynamic range of the signal.

[1]The amount of light entering the camera sensor.

- This noise originates from reading the signal from the sensor, in this case through the field effect transistor (FET) of a CCD chip.
- The general form of the power spectral density of readout noise is:

$$S_{\mathrm{nm}}(\omega) \approx \begin{cases} \omega^{-\beta} & \omega < \omega_{\mathrm{min}} \quad \beta > 0, \\ k & \omega_{\mathrm{min}} < \omega < \omega_{\mathrm{max}}, \\ \omega^{\alpha} & \omega > \omega_{\mathrm{max}} \quad \alpha > 0. \end{cases}$$

where α, β are constants, and ω is the radial frequency at which the signal is transferred.

- At very low readout rates ($\omega < \omega_{\mathrm{min}}$) the noise is pink.
- Readout noise can be reduced to manageable levels by appropriate readout rates and proper electronics. At very low signal levels, readout noise can still become a significant component in the overall SNR.

Noise

- In the case of CCD sensors, all the pixels in a sensor pass through a common architecture and are essentially subject to the same sources of noise during the readout process.
- Therefore, the read noise of a CCD sensor can be described by a single readout noise value.
- The readout process for sCMOS however is different, sCMOS sensors are often referred to as Active Pixel Sensors (APS) since each pixel has its own amplifier circuit.

Factors

- Read noise is inherent to the readout process of the sensor itself,
  - but cameras from different manufacturers that are based around the same sensor can have some significant differences in how the sensor has been implemented.
- sCMOS cameras are remarkably flexible imaging devices and have several settings that allow them to be optimised for different applications such as high speed, or for high dynamic range imaging.
- Different cameras will have different behaviours as settings are adjusted and this also includes how read noise is affected.

The process of a CMOS camera is as follows:

- Photons hit the sensor and generate charge (electrons),
- The photo-generated charge is converted to an analog voltage for each pixel amplifier,
- These pixel voltages are transferred to the column bus via a row select signal,
- The analog voltages are then converted to digital signals via columns of analog to digital (A/D) converters,
- The final digitised signals are then read out sequentially at a pixel readout speed which normally can be set at different speeds,

- Noise associated with the gate capacitor of an FET (Field Effect Transistor) is known as **Johnson-Nyquist** noise and can be non-negligible.
- The output RMS value of the noise voltage (v_{JN}) is given by:

$$v_{\text{JN}} = \sqrt{kT/C} \quad \text{V}.$$

where C is gate switch capacitance, k_{B} is Boltzmann's constant, and T is the absolute temperature of the CCD chip in K.

1 fF	2 mV	10 pF	20 µV
10 fF	640 µV	100 pF	6,4 µV
100 fF	200 µV	1 nF	2 µV

Table 6: Effect on the noise on the capacitor value.

- This happens in the amplitude quantization, occurring in the ADC[2].
- The noise is additive and independent of the signal when the number of levels (L) is less than 2^4 where $B = 4$.
- When a signal is converted to electricity, it has a minimum and maximum electrical value, due to quantisation.
- If the ADC is adjusted so that 0 corresponds to the minimum electrical value and $2B - 1$ is the maximum value, SNR is:

$$SNR = 6B + 11 \quad \text{dB}$$

- For $B \geq 8$ bits, this means a SNR ≥ 59 dB.
- This can usually be ignored as the total SNR is typically dominated by the smallest SNR. In CCD cameras this is photon noise.

[2] Analog to Digital Converter

Noise

- Perlin noise is a type of gradient noise developed by Ken Perlin in 1983.
- The need arose from the machine-looking textures used in CGI in the 80's.
- It has many uses, including but not limited to:
  - procedure generating terrain, applying pseudo-random changes to a variable, and assisting in the creation of image textures.
- It is most commonly implemented in two, three, or four dimensions, but can be defined for any number of dimensions.
- Used frequently to generate textures with extremely limited memory.
- Similar methods exists such as fractal noise and simplex noise

- First thing is to load all the necessary modules.
- The only novel module you may have not encountered is `noise`
  - A library including native-code implementations of Perlin “improved” noise and Perlin simplex noise.

C.R. 14

python

```python
import noise
import numpy as np
from PIL import Image
import matplotlib.pyplot as plt
```

- The only module worth mentioning is `PIL` which is an image library for python.

- Let's have a look at the parameters we can play with:

C.R. 15

python

```python
shape = (256,256)
scale = 200
octaves = 6
persistence = 0.5
lacunarity = 2.0
seed = 51
```

shape Dimensions of the image,

scale altitude in which to see the noise

octaves number of layers of the algorithm

persistence how much more each successive value brings

lacunarity level of detail per pass

seed the initial value for the RNG.

Below is the main function for the perlin noise generation.

C.R. 16

python

```python
world = np.zeros(shape)
for i in range(shape[0]):
    for j in range(shape[1]):
        world[i][j] = noise.pnoise2(i/scale,j/scale,
                                    octaves=octaves,
                                    persistence=persistence, lacunarity=lacunarity,
                                    repeatx=1024, repeaty=1024,
                                    base=seed)
```

The above function generate the noise from calling the `pnoise2` function from the `noise` library and writes the results to the numpy array `world`.

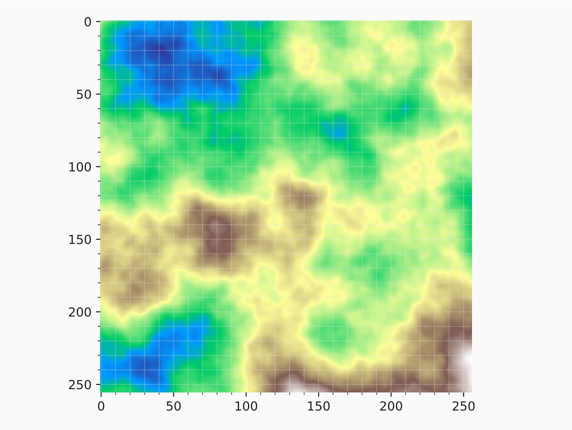


Figure 88: 2D Perlin Noise plot. To give the aesthetics of a map, a colour map `cmap='terrain'` was used

- While this map is quite good we can make it more pop by introducing the height to the image to create a 3D plot.
- For plotting this in 3 dimensions, we must initialise 2 more arrays which will contain the x-y co-ordinates of our world.

C.R. 17

python

```python
from mpl_toolkits.mplot3d import axes3d
```

C.R. 18

python

```python
lin_x = np.linspace(0,1,shape[0],endpoint=False)
lin_y = np.linspace(0,1,shape[1],endpoint=False)
x,y = np.meshgrid(lin_x,lin_y)
```

- Now we only need to write the code which will plot this 3D data.

C.R. 19

python

```python
fig = plt.figure()
ax = fig.add_subplot(projection='3d')
ax.plot_surface(x,y,world,cmap='terrain')

for spine in ax.spines.values():
    spine.set_visible(False)

cp.store_fig("perlin-plot-3d-map",
             close = True)
```

- We use `cmap = 'terrain'` to make it look like a topological map.
- We also disable the axis with `spine.set_visible(False)`.

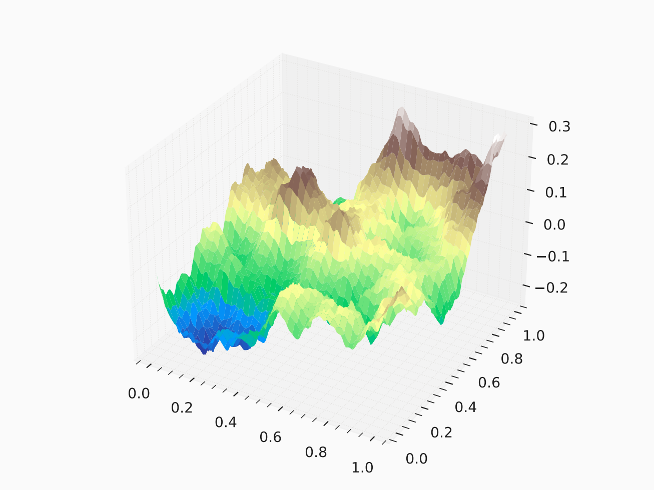


Figure 89: Our newly generated 3D terrain based on Perlin noise.

CCD Charge Coupled Device. 126

[1] Michael F Deering. **"The limits of human vision".** In: *2nd international immersive projection technology workshop*. Vol. 2. 1. 1998.

[2] Hawesthoughts. ***Colorblindness Wheel.*** 2023. URL: https://commons.wikimedia.org/wiki/File:Color_blindness_wheels.svg.

[3] Colour Blind Awareness. ***Types of Colour Blindness.*** 2024. URL: https://www.colourblindawareness.org/colour-blindness/types-of-colour-blindness/.

[4] Colour Blind Awareness. ***Inherited Colour Vision Deficiency.*** 2024. URL: https://www.colourblindawareness.org/colour-blindness/causes-of-colour-blindness/inherited-colour-vision-deficiency/.

[5] Amanda Jefferies et al. **"Do You See What I See?: Understanding the Challenges of Colour-Blindness in Online Learning".** In: *Proceedings of 10th European Conference for E-Learning*. Academic Conferences Ltd. 2011.

[6] Susanne Kohl et al. **"Achromatopsia".** In: (2018).

[7] American Optometric Association. ***Color vision deficiency***. 2024. URL: https://www.aoa.org/healthy-eyes/eye-and-vision-conditions/color-vision-deficiency?sso=y.

[8] Gigahertz-Optik. ***Spectral Sensitivity of the Human Eye***. 2002. URL: https://www.gigahertz-optik.com/en-us/service-and-support/knowledge-base/basics-light-measurement/light-color/spectr-sens-eye/.

[9] Hyperphysics. ***The Color-Sensitive Cones***. 2024. URL: http://hyperphysics.phy-astr.gsu.edu/hbase/vision/colcon.html.

[10] Zoltan Derzsi and Robert Volcic. **"Not only perception but also grasping actions can obey Weber's law"**. In: *Cognition* 237 (2023), p. 105465.

[11] Edward H Adelson. **"Checkershadow illusion. 1995"**. In: *URL http://web.mit. edu/persci/people/adelson/checkershadow_illusion. html* (2005).

[12] BBC. ***Optical illusion: Dress colour debate goes global***. 2015. URL: https://www.bbc.com/news/uk-scotland-highlands-islands-31656935.

[13] PolBr. ***sRGB colors situated at calculated position in CIE 1931 chromaticity diagram (edited from CIExy1931_sRGB.svg). Luminance Y set so that R+G+B = 1 to avoid bright lines toward primaries' complementary colours.*** 2021. URL: https://commons.wikimedia.org/wiki/File:SRGB_chromaticity_CIE1931.svg.

[14] ***Multimedia systems and equipment - Colour measurement and management - Part 2-1: Colour management - Default RGB colour space - sRGB.*** Standard. Geneva, CH: International Electrotechnical Commision, 1999.

[15] R. Rehák, P. Bodrogi, and J. Schanda. **"On the use of the sRGB colour space".** In: *Displays* 20.4 (1999), pp. 165–170. ISSN: 0141-9382. DOI: https://doi.org/10.1016/S0141-9382(99)00019-0. URL: https://www.sciencedirect.com/science/article/pii/S0141938299000190.

[16] Danny Pascale. **"A review of rgb color spaces... from xyy to r'g'b"'.** In: *Babel Color* 18 (2003), pp. 136–152.

[17] Andrew Rodney. **"The role of working spaces in Adobe applications".** In: *Technical Paper. Adobe. Retrieved* (2008), pp. 05–09.

[18] ***Photography and graphic technology – Extended colour encodings for digital image storage, manipulation and interchange.*** Standard. Geneva, CH: International Organization for Standardization, 2013.

[19] Kevin E Spaulding, Geoffrey J Woolfe, and Edward J Giorgianni. **"Reference input/output medium metric rgb color encodings".** In: *Color Imaging Conference*. Vol. 1. 2000, p. 2.

[20] ***Multimedia systems and equipment - Colour measurement and management - Part 2-5: Colour management - Optional RGB colour space - opRGB.*** Standard. Geneva, CH: International Electrotechnical Commision, 2007.

[21] Inc. Reproductions. ***Understanding the Differences Between RGB and CYMK Colors.*** 2024. URL: https://reproductionsinc.com/understanding-the-differences-between-rgb-and-cymk-colors/.

[22] Slippens. ***Three examples of color halftoning with CMYK separations***. 2009. URL: https://en.wikipedia.org/wiki/File:Halftoningcolor.svg.

[23] S. A. Eugster. ***CbCr plane of the YCbCr color space, with a Y value of 0.5, Cb on the horizontal x axis going from -1 to 1, Cr on the y axis going from -1 to 1 as well.*** 2010. URL: https://commons.wikimedia.org/wiki/File:YCbCr-CbCr_Scaled_Y50.png.

[24] J. Sneyers. ***Various image formats are categorized by scope in this diagram, along two axes: Raster vs Vector, and Authoring vs Delivery.*** 2022. URL: https://commons.wikimedia.org/wiki/File:Image_formats_by_scope.svg.

[25] Adobe. ***Lossy vs Lossless Compression Differences and When to Use***. 2024. URL: https://www.adobe.com/uk/creativecloud/photography/discover/lossy-vs-lossless.html.

[26] Khalid Sayood. ***Introduction to data compression***. Morgan Kaufmann, 2017.

[27] Alistair Moffat. **"Huffman coding"**. In: *ACM Computing Surveys (CSUR)* 52.4 (2019), pp. 1–35.

[28] Michael Gabler. ***Various image formats are categorized by scope in this diagram, along two axes: Raster vs Vector, and Authoring vs Delivery.*** 2011. URL: https://commons.wikimedia.org/wiki/File:Felis_silvestris_silvestris_small_gradual_decrease_of_quality.png.

[29] Mark Stanford Robbins. **"Electron-multiplying charge coupled devices–emccds"**. In: *Single-Photon Imaging* (2011), pp. 103–121.

[30] Bedabrata Pain et al. **"A back-illuminated megapixel CMOS image sensor"**. In: (2005).

[31] Estrin. ***Kodak's First Digital Moment***. 2015. URL: https://archive.nytimes.com/lens.blogs.nytimes.com/2015/08/12/kodaks-first-digital-moment/.

[32] Teledyne. ***Types of Camera Sensor***. 2024. URL: https://www.photometrics.com/learn/camera-basics/types-of-camera-sensor.

[33] Eric R Fossum and Donald B Hondongwa. **"A review of the pinned photodiode for CCD and CMOS image sensors".** In: *IEEE Journal of the electron devices society* (2014).

[34] Eric R Fossum. **"Active pixel sensors: Are CCDs dinosaurs?"** In: *Charge-Coupled Devices and Solid State Optical Sensors III*. Vol. 1900. SPIE. 1993, pp. 2–14.

[35] Filya1. ***CMOS image sensor.*** 2009. URL: https://en.m.wikipedia.org/wiki/File:Matrixw.jpg.

[36] Olympus. ***Introduction to CMOS Image Sensors.*** 2024. URL: https://www.olympus-lifescience.com/en/microscope-resource/primer/digitalimaging/cmosimagesensors/.

[37] Eric Fox. ***CMOS TDI: A Game Changer in High-Fidelity Imaging.*** 2019. URL: https://possibility.teledyneimaging.com/cmos-tdi-a-game-changer-in-high-fidelity-imaging/.

[38] Junichi Nakamura. ***Image sensors and signal processing for digital still cameras.*** CRC press, 2017.

[39] Harris R Miller. **"Color filter array for CCD and CMOS image sensors using a chemically amplified thermally cured pre-dyed positive-tone photoresist for 365-nm lithography".** In: *Advances in Resist Technology and Processing XVI*. Vol. 3678. SPIE. 1999, pp. 1083–1090.

[40] Ralph Dammel. ***Diazonaphthoquinone-based resists.*** Vol. 11. SPIE press, 1993.

[41] Cmglee. ***A Bayer pattern on a sensor in isometric perspective/projection.*** 2020. URL: https://commons.wikimedia.org/wiki/File:Bayer_pattern_on_sensor.svg.

[42] FlickreviewR. ***Sony F828 Camera.*** 2009. URL: https://commons.wikimedia.org/wiki/File:Sony_F828.jpg.

[43] teledyneimage. ***Linearity: Image Topics.*** 2024. URL: https://www.photometrics.com/learn/imaging-topics/linearity.

[44] Teledyne. ***How to Evaluate Camera Sensitivity.*** 2021. URL: https://www.flir.com/discover/iis/machine-vision/how-to-evaluate-camera-sensitivity/.

[45] TeledyneCurrent. ***Dark Current.*** 2024. URL: https://www.photometrics.com/learn/advanced-imaging/dark-current.

[46] Crisp. ***Getting out of Auto: Understanding ISO on your digital camera.*** 2016. URL: https://newatlas.com/photography-iso-range-setting-guide/44599/.

[47] EdmundOptics. ***Anatomy of a Lens.*** 2024. URL: https://www.edmundoptics.com/knowledge-center/application-notes/imaging/the-anatomy-of-a-lens/.

[48] DarkRoom. ***Effects of Focal Length and Angle.*** 2024. URL: https://thedarkroom.com/focal-length/.

[49] Morio. ***Fujinon XF100-400mm F4.5-5.6 R LM OIS WR cutaway.*** 2016. URL: https://commons.wikimedia.org/wiki/File:Fujinon_XF100-400mm_F4.5-5.6_R_LM_OIS_WR_cutaway_2016_China_P%26E.jpg.

[50] DrBob. ***Simple zoom lens.*** 2005. URL: https://commons.wikimedia.org/wiki/File:Zoomlens1.svg.

[51] Zurek. ***Chromatic aberration.*** 2006. URL: https://commons.wikimedia.org/wiki/File:Chromatic_aberration_(comparison).jpg.

[52] European Broadcasting Union. **“Enhanced 625-line Phased Alternate Line (PAL) television”.** In: (1997).

[53] Amer Al-Rahayfeh and Miad Faezipour. **“Enhanced frame rate for real-time eye tracking using circular hough transform”.** In: *2013 IEEE Long Island Systems, Applications and Technology Conference (LISAT)*. IEEE. 2013, pp. 1–6.

[54] Mary C Potter et al. **“Detecting meaning in RSVP at 13 ms per picture”.** In: *Attention, Perception, & Psychophysics* 76 (2014), pp. 270–279.

[55] Benjamin Tag et al. **“In the eye of the beholder: The impact of frame rate on human eye blink”.** In: *Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems*. 2016, pp. 2321–2327.

[56] Rudolf F Graf. ***Modern dictionary of electronics.*** Elsevier, 1999.

[57] Ji-Won Lee et al. **"Screenshot identification by analysis of directional inequality of interlaced video"**. In: *EURASIP Journal on Image and Video Processing* 2012 (2012), pp. 1–15.

[58] ***Interlaced Video & Deinterlacing for streaming***. 2021. URL: https://blog.video.ibm.com/streaming-video-tips/interlaced-video-deinterlacing-for-streaming/.

[59] Logan. ***A monochrome CRT as seen inside a Macintosh Plus computer***. 2018. URL: https://commons.wikimedia.org/wiki/File:Macintosh_Plus_interior.jpg.

[60] Expromo. ***Led Display Facts***. 2020. URL: https://www.expromo.eu/en/led-display-facts/.

[61] ***LCD vs. LED: What is LCD & LED***. 2023. URL: https://www.linsnled.com/difference-between-lcd-and-led.html.

[62] ***Philips Test Card***. 2008. URL: https://commons.wikimedia.org/wiki/File:Philips_PM5544.svg.

[63] Logical Increments. ***Information About Screen Resolution***. 2023. URL: https://www.logicalincrements.com/resolution.

[64] Genelec. ***Audio Test Signals.*** 2024. URL: https://www.genelec.com/audio-test-signals.