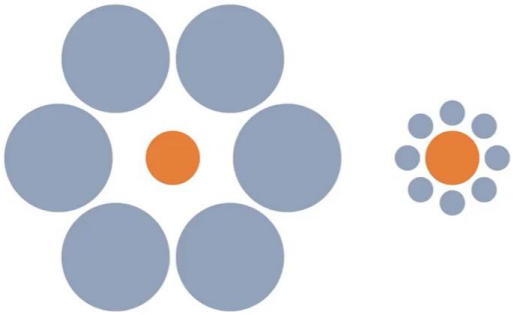
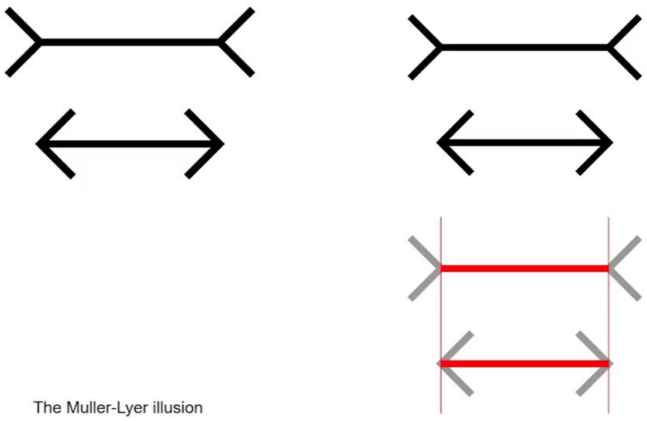
What is Computer Vision?

- The objective of computer vision is to make computers see and interpret the world like humans and possibly even better than us.
- It involves the development of a theoretical and algorithmic basis to achieve automatic visual understanding.
- It is an interdisciplinary field that works for high-level understanding from digital images or videos.
- Computer vision has a dual goal.
    - From the biological science point of view, computer vision aims to come up with computational models for human visual system.
    - From the engineering point of view, computer vision aims to build autonomous systems to perform some of the tasks which the human visual system can perform and even surpass it in many cases.
- Computer vision is transitioning from a nascent stage and is proving to be incredibly useful in several application areas. It's in our smart phone cameras which are able to recognize faces and smiles. It's in self driving cars, reading traffic signs and watching for pedestrians. It's in factory robots, monitoring for problems and navigating around human coworkers. This is just the beginning.

Human visual system has well-known set of shortcomings that are experienced as optical illusions, where we tend to misjudge size, color, and movement.



The Ebbinghaus illusion

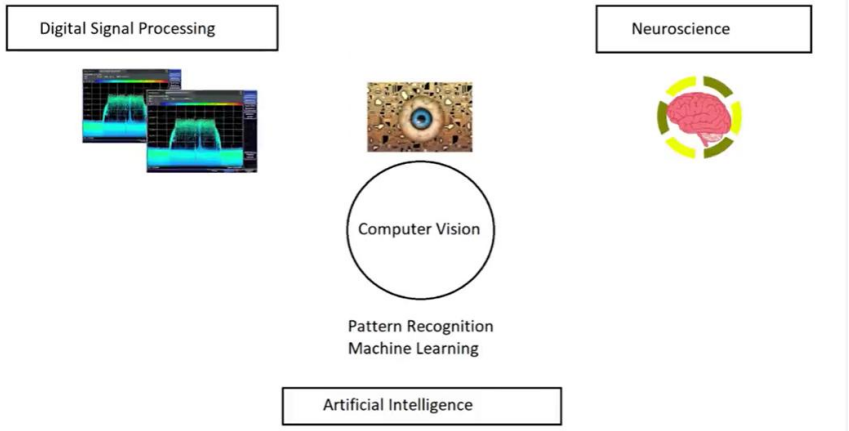


The Muller-Lyer illusion

- ❖ The physical world is 3D, and human vision perceives depth using two eyes (stereo vision) to form a rich 3D understanding. In contrast, most current computer vision algorithms, like YOLO, work on 2D images and lack true depth perception. However, monocular depth estimation aims to make CV more like human vision by enabling single-camera systems to estimate depth and understand 3D structure from flat images.

Related Fields of Computer Vision:

The field of computer vision heavily incorporates concepts from the areas of digital signal processing, neuroscience, and artificial intelligence. Artificial intelligence and computer vision share topics such as pattern recognition and machine learning.



Timelines and Milestones:

- To understand where we are today in the field of computer vision, it is important to know the history and the key milestones that shaped this rapidly evolving field.
- The field of computer vision emerged in the 1950s with research along three distinct lines. **Replicating the eye, replicating the visual cortex and replicating the rest of the brain.** These are cited in order according to their level of difficulty. One early breakthrough came in 1957 in the form of the Perceptron machine. The giant machine tightly tangled with wires was the invention of psychologist and computer vision pioneer, Frank Rosenblatt. Then the same year marked the birth of the pixel. In the spring of 1957, the National Bureau of Standards Scientists, Russell Kirsch, took a photo of his son Walden and scanned it into the computer. To fit the image into computer's limited memory, he divided the picture into a grid. This five centimeters square photo was the first digital image ever created.
- In 2014, a team of researchers at the University of Montreal introduced the idea that machines can learn faster by having two neural networks compete against each other. One network attempts to generate fake data that looks like the real thing and the other network tries to discriminate the fake from the real. Over time, both networks improve. The generator produces data so real that the discriminator can't tell the difference. Generative

adversarial networks are considered a significant breakthrough in computer vision in the past few years.

Computer Vision Applications:

- There has also been great progress in traditional application areas like **multimedia**, **robotics**, and **medical imaging**. Moreover, new application areas keep arising such as **augmented reality**, **autonomous driving**, **Internet of things**, **human-computer interaction** and **vision for the blind**.

- The demand for fast and robust processing of vision task is so high that **vision processing units** are emerging as a new class of processors to complement CPUs and GPUs which are Graphics Processing Units.
    - **Visual Surveillance**, **biometrics application (Iris Recognition)**
    - **Seeing AI** is a Microsoft research project for people with visual impairments. **Lookout** by Google.

Light Sources:

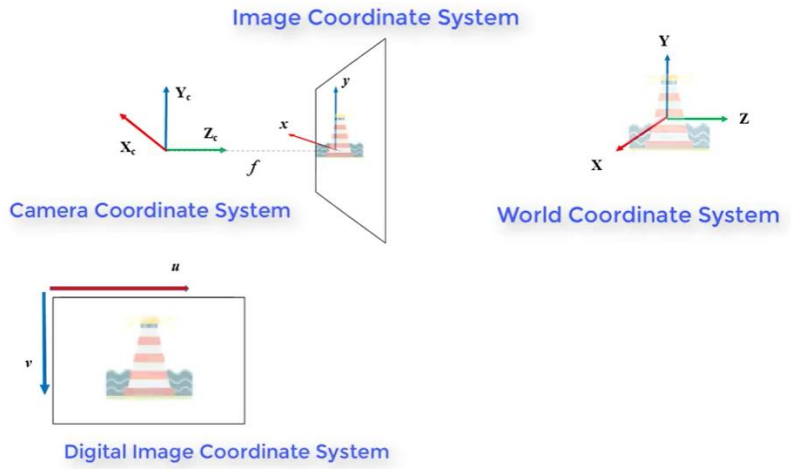
Images cannot exist without light. To produce an image, the scene must be illuminated with one or more light sources.

In general, the word light refers to the visible spectrum of the electromagnetic radiation. Visible light spans to wavelengths between 400-700 nanometers which is a very minute fraction of the entire electromagnetic spectrum.

- Machine vision, which deals with vision tasks and controlled setting capitalize on the principles of lighting.

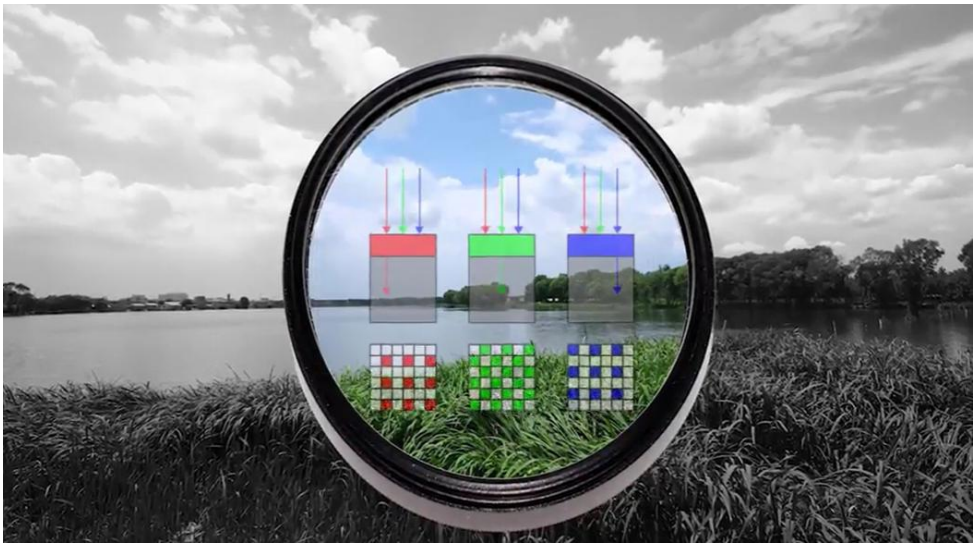
Pinhole Camera Model:

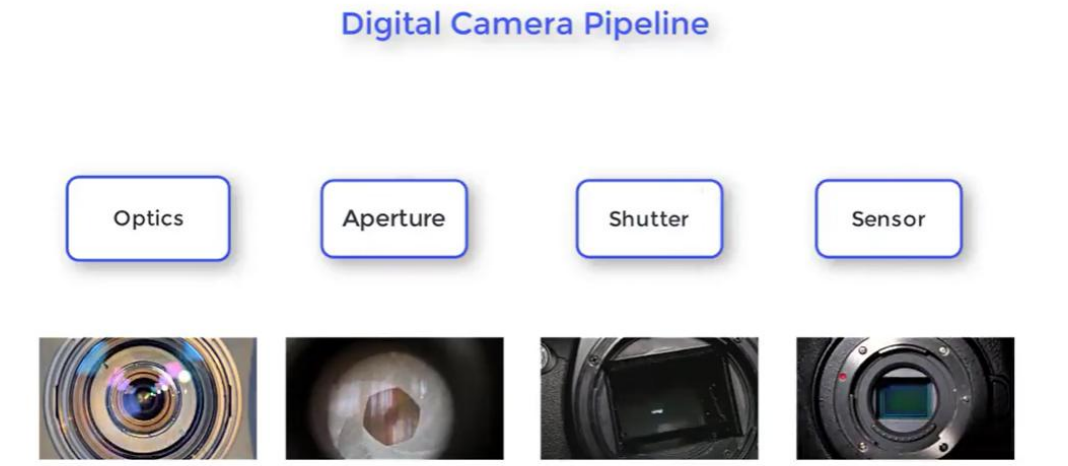
The counterpart for the human eye in computer vision is a camera. There are many variants of a camera which differ in the lens configurations and image sensors. The most basic camera model is a pinhole camera model.

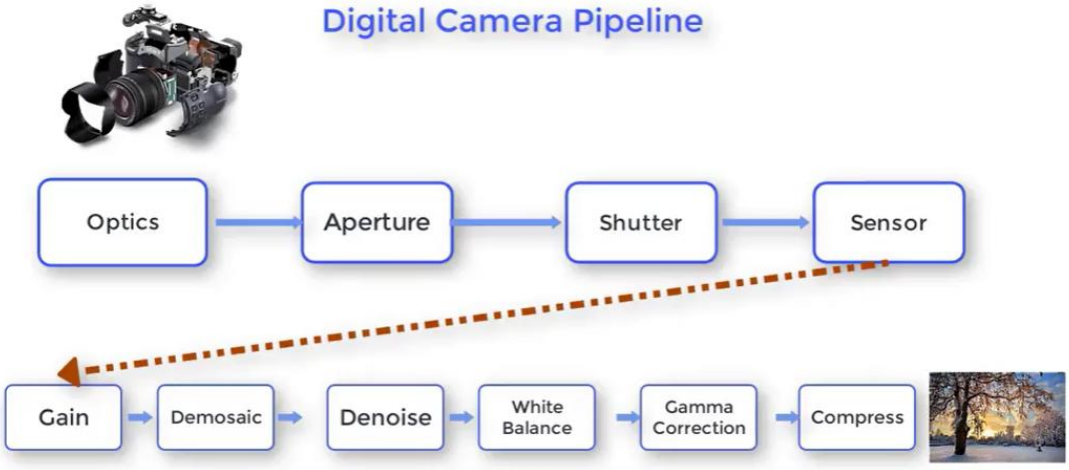


Digital Camera:

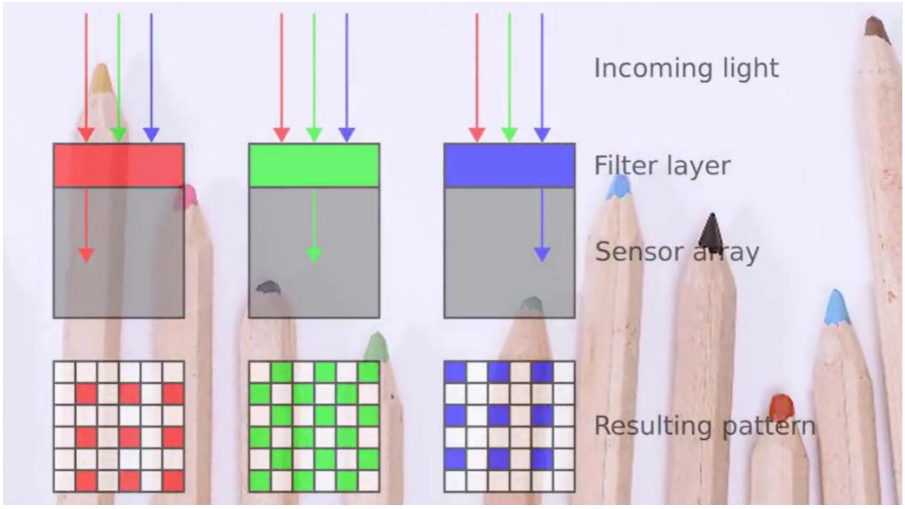
- Re-inventing the eye is the area where we've had the most success in the field of computer vision.
- How are the photos from the light sources arriving in the sensor converted into the digital RGB values that we observe in the digital image?



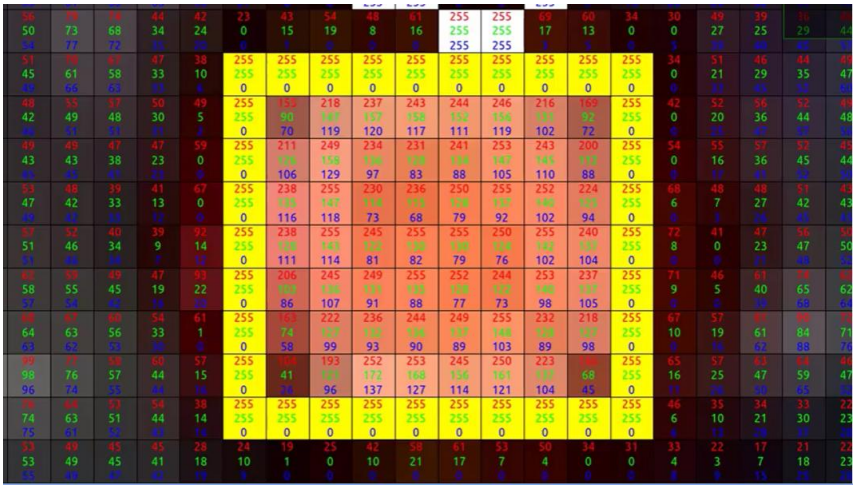




Color Theory:



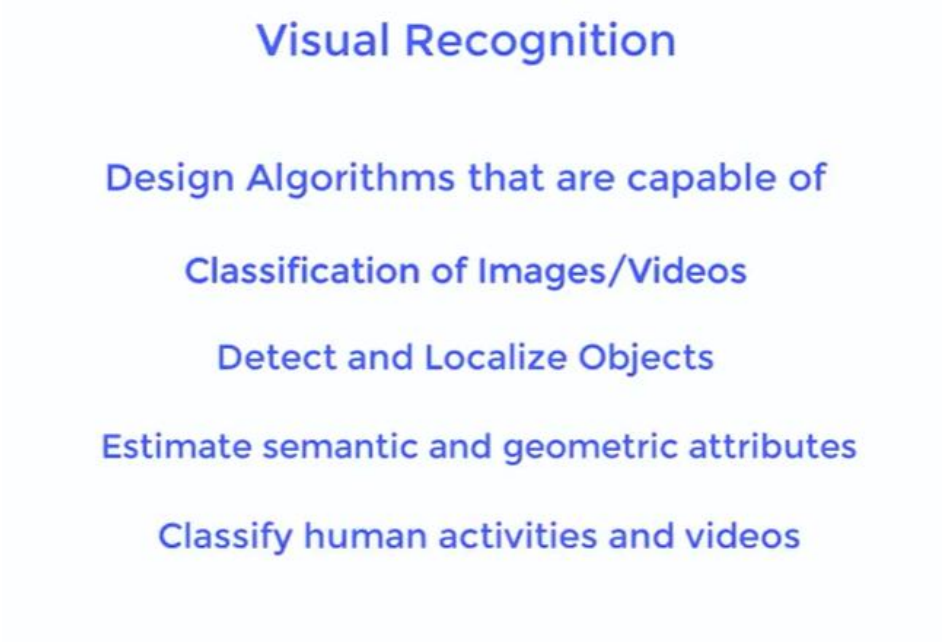
Each pixel location in an RGB color image contains three intensity values.



David Marr's Framework and Early Vision Concepts

One of the first paradigms of computer vision was proposed by David Marr. He described a general framework for understanding visual perception and touched on broader questions about how the brain and its functions can be studied and understood.

- One of the important contribution made by Marr is the representational framework for vision using primal sketch. Primal sketch is similar to a pencil sketch drawn quickly by an artist as an impression.
- An application which create the color distribution of picture. (3D)



Mathematics for Computer Vision

Computer Vision is used to solve vital problems in a vast array of fields including medical imaging, surveillance, face and object detection and identification. The techniques that *linear algebra* provides for solving complicated mathematical models are essential to solve problems in each of these fields.

- **Singular value decomposition (SVD)** is the most common and useful linear algebra technique in Computer Vision because it helps to achieve the goal of Computer Vision, which is to explain the three dimensional world through two dimensional pictures.

Calculus is the mathematical study of continuous change, in the same way that geometry is the study of shape and algebra is the study of generalizations of mathematic operations.

- Calculus has two major branches, differential calculus, concerning instantaneous rates of change and slopes of curves. And then we have integral calculus.
- **Integral calculus** - concerned with the theory and applications of integrals. It deals with total size or value, such as lengths, areas, and volumes.
- Computer Vision uses derivatives, integrals and partial differential equations extensively in several low and mid-level vision tasks.

Image processing focuses on improving images, adjusting brightness, sharpening details, and reducing noise, without necessarily understanding what's in them.

Computer vision, on the other hand, takes things further by enabling machines to recognize and interpret images and videos as humans do. This makes tasks like facial recognition, object detection, and real-time scene analysis possible.

Thank You!