Computer Vision and Pattern Recognition - COE 406

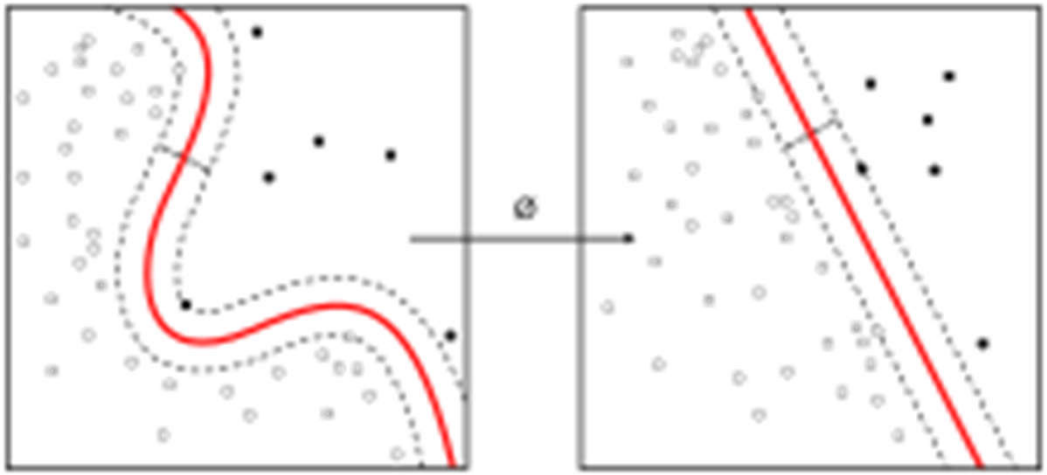
Department of Computer Engineering
Faculty of Engineering
University of Baghdad
2020/2021

Today

- Introduction to Computer vision and Pattern Recognition
- Overview
- Requirements

- Computer vision is the ability of computers to see (To bridge the gap between pixels and meaning)
  - Machine Vision
  - Robot Vision
  - Image Understanding
  - Video Analysis

- Pattern recognition is the automated recognition of patterns and regularities in data.
  - The objective behind pattern recognition algorithms is to provide a reasonable answer for all possible data and to classify input data into objects or classes based on certain features



- ***The problem of computer vision appears simple because it is trivially solved by people, even very young children.*** Nevertheless, it largely remains an unsolved problem based both on the limited understanding of biological vision and because of the complexity of vision perception in a dynamic and nearly infinitely varying physical world.

- Smartphones have cameras, and taking a photo or video and sharing it has never been easier, **resulting** in the **incredible growth** of modern **social networks** like Instagram.

- **YouTube** might be the **second largest search engine** and **hundreds of hours of video** are **uploaded every minute** and **billions of videos are watched every day.**

- The internet is comprised of text and images. It is relatively straightforward to index and search text, but in order to index and search images, algorithms need to know what the images contain. For the longest time, the content of images and video has remained opaque, best described using the meta descriptions provided by the person that uploaded them.

- To get the most out of image data, we need computers to “see” an image and understand the content.

- Image processing is the process of creating a new image from an existing image, typically simplifying or enhancing the content in some way. It is a type of digital signal processing and is not concerned with understanding the content of an image.

- A given computer vision system may require **image processing** to be applied to raw input, e.g. **pre-processing images**.

- Examples of image processing include:
  - Normalizing photometric properties of the image, such as **brightness** or **color**.
  - Cropping the bounds of the image, such as **centering an object** in a photograph.
  - **Removing digital noise** from an image, such as digital artifacts from low light levels.

- The goal of computer vision is to extract useful information from images. This has proved a surprisingly challenging task; it has occupied thousands of intelligent and creative minds over the last four decades, and despite this we are still far from being able to build a general-purpose “seeing machine.”

— Page 16, Computer Vision: Models, Learning, and Inference, 2012.

Vision is

- Misleadingly easy
- Deceptive
- Computationally demanding
- Critical to many applications

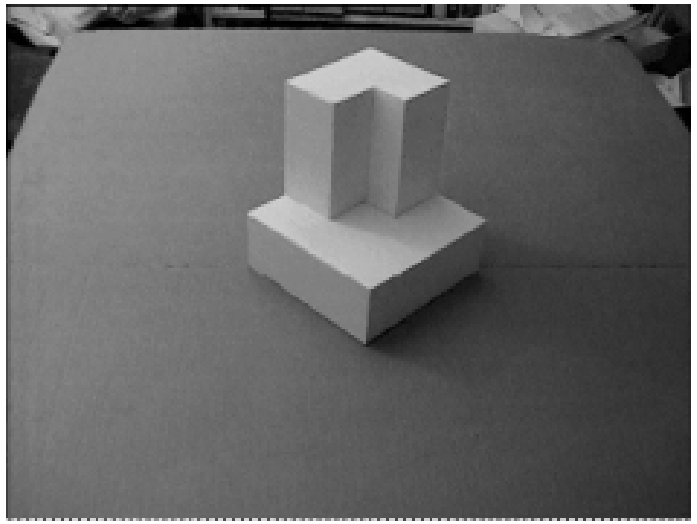
Vision is deceivingly easy

- We see effortlessly
  - seeing seems simpler than “thinking”
  - we can all “see” but only select gifted people can solve “hard” problems like chess
  - we use nearly 70% of our brains for visual perception!
- All “creatures” see
  - frogs “see”
  - birds “see”
  - snakes “see”

but they do not see alike

Vision is deceivingly easy

- The M.I.T. summer vision program
  - summer of 1965
  - point TV camera at stack of blocks
  - locate individual blocks
    - recognize them from small database of blocks
  - describe physical structure of the scene
    - support relationships
- Formally ended in 1985

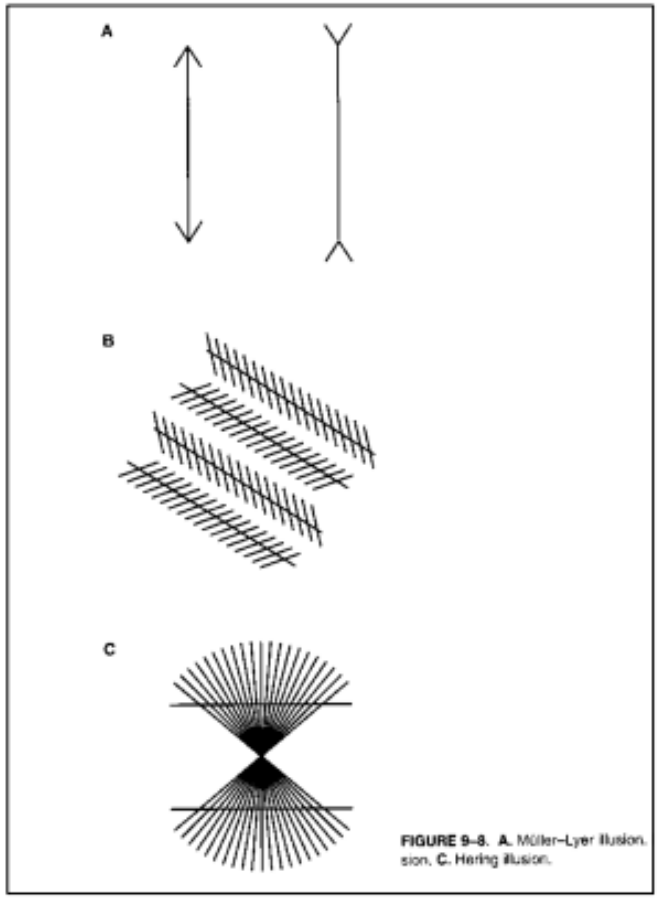


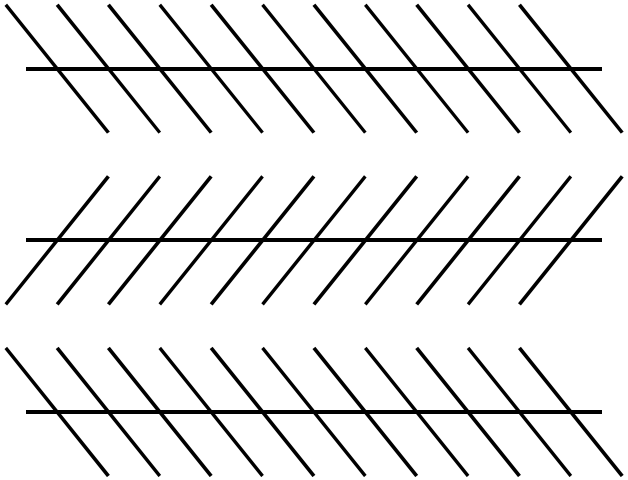
Vision is deceptive

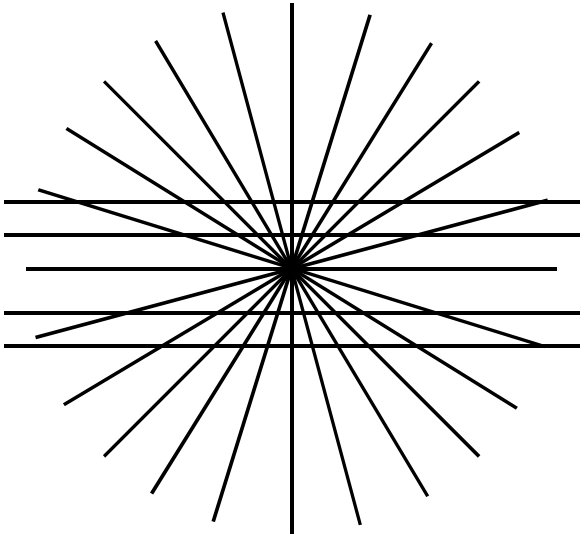
- Vision is an exceptionally strong sensation
  - vision is immediate
  - we perceive the visual world as external to ourselves, but it is a reconstruction within our brains
  - we regard how we see as reflecting the world “as it is;” but human vision is
    - subject to illusions
    - quantitatively imprecise
    - limited to a narrow range of frequencies of radiation
    - passive

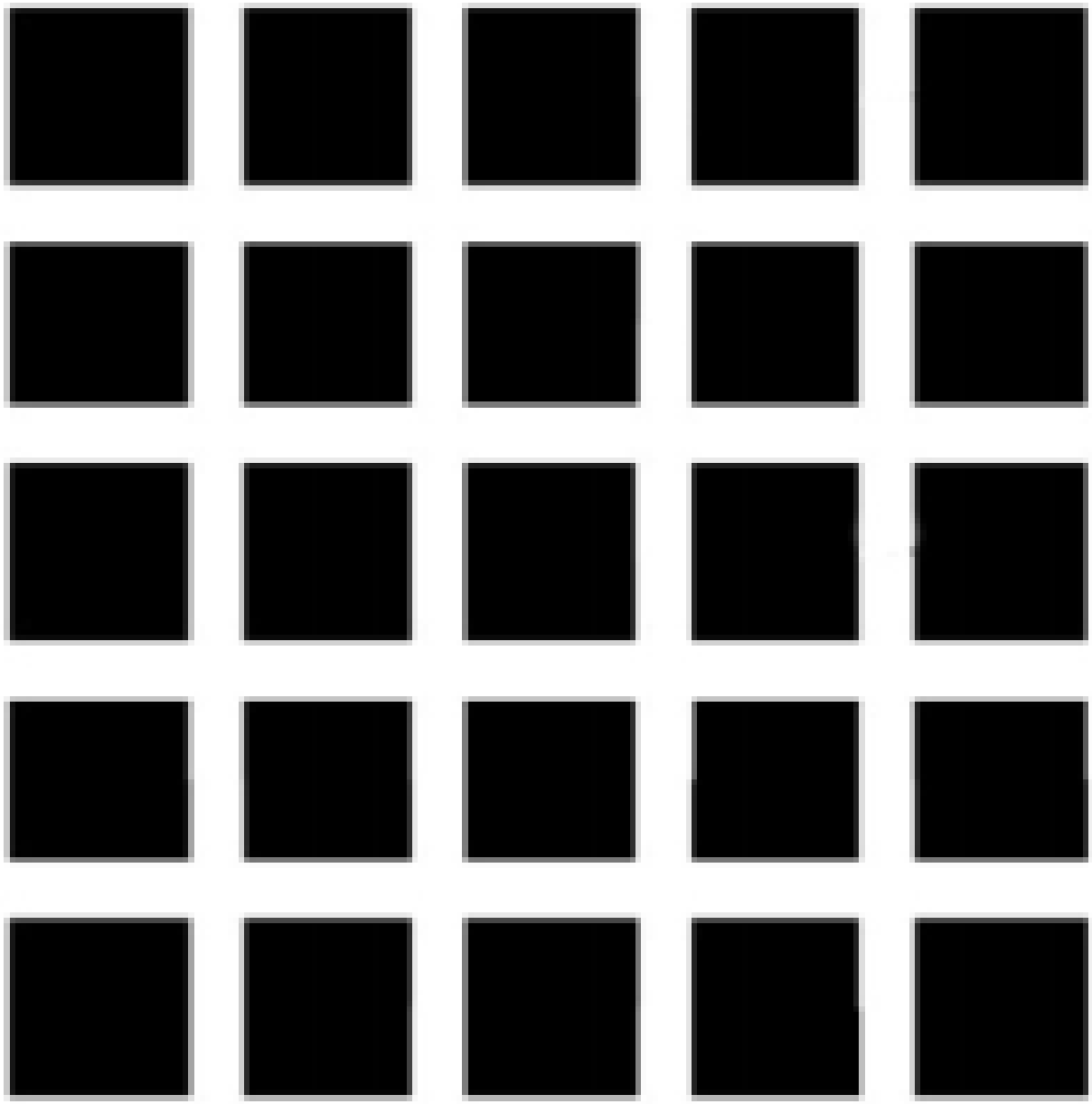
Vision is deceptive

- Human vision is
  - subject to illusions
  - quantitatively imprecise
  - limited to a narrow range of frequencies of radiation
  - passive

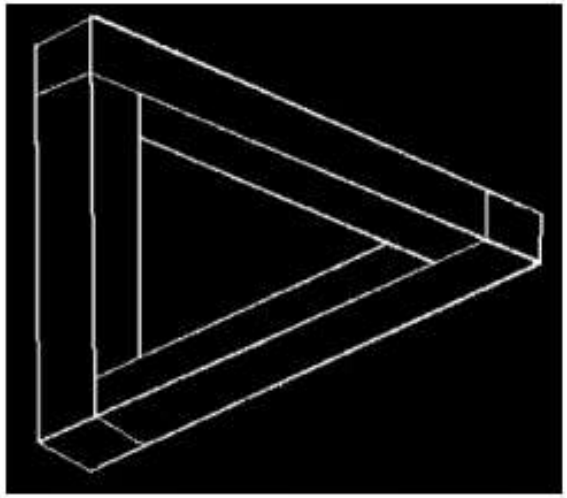


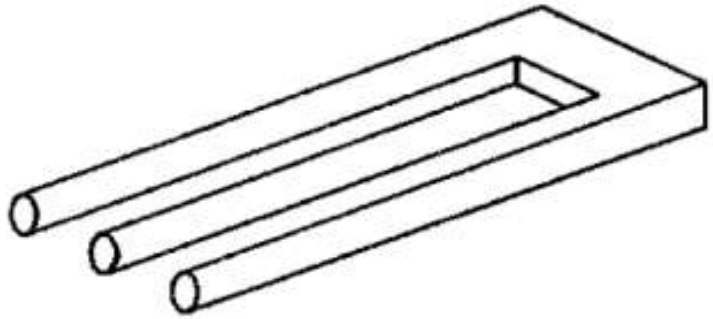






More illusion

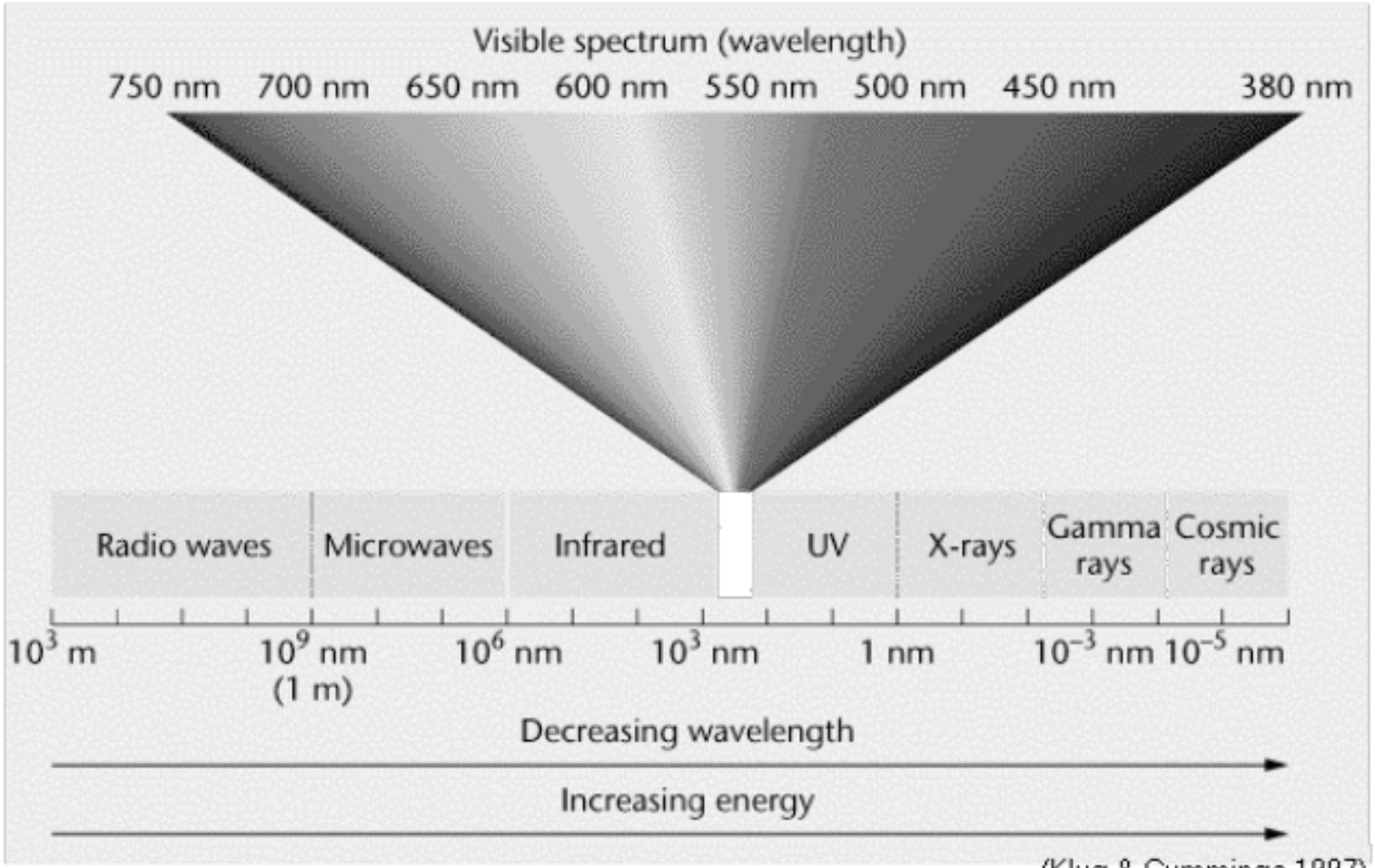




- We can see impossible figures

Spectral limitations of human vision

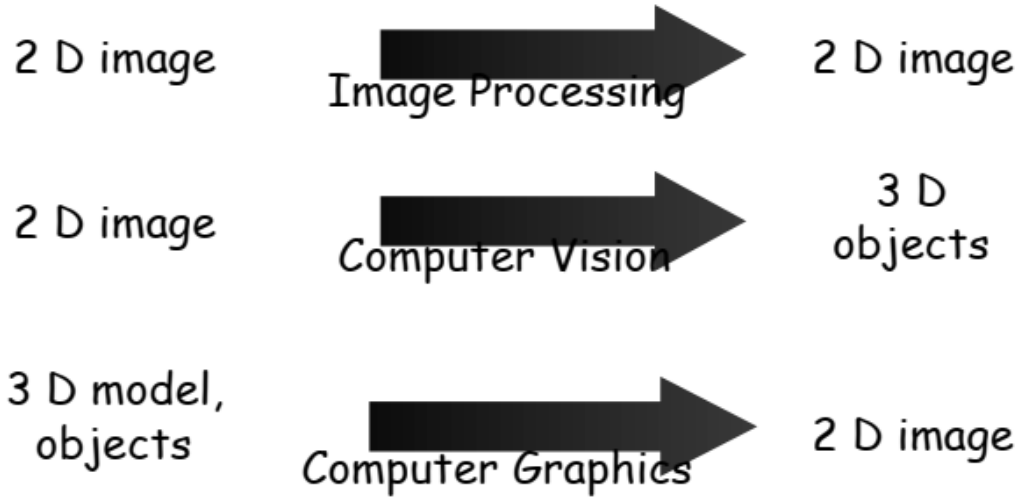
- We “see” only a small part of the energy spectrum of sunlight
  - we don’t see ultraviolet or lower frequencies of light
  - we don’t see infrared or higher frequencies of light
  - we see less than .1% of the energy that reaches our eyes
- But objects in the world reflect and emit energy in these and other parts of the spectrum



(Klug & Cummings 1997)

Non-human vision

- Infrared vision
- Polarization vision
  - navigation for birds and insects
- Ultrasound vision
- X-ray vision!
- RADAR vision



Related Fields: AI, pattern recognition, machine learning, signal processing, neural networks, cognitive vision.

Generic Object Recognition

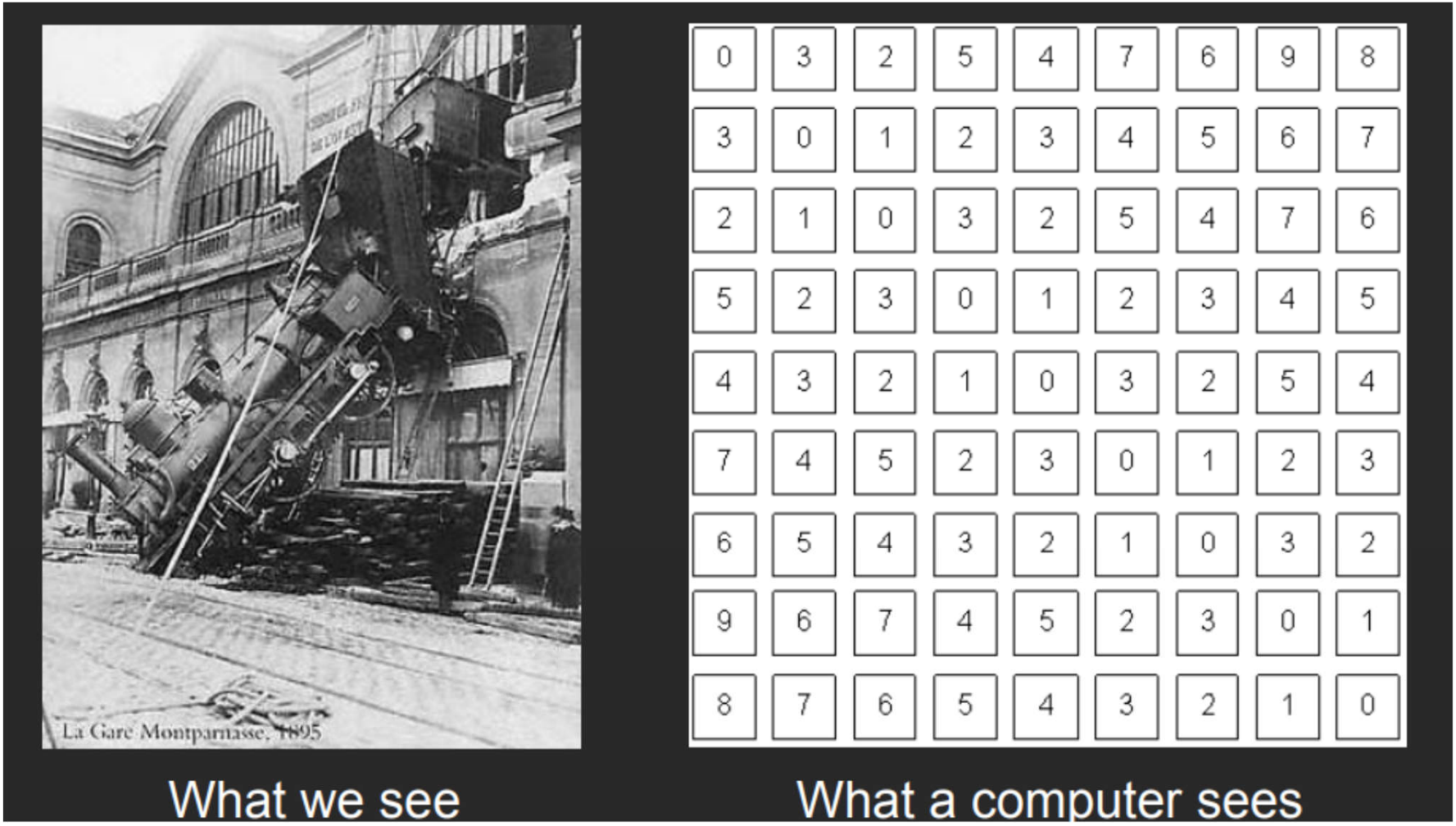







- Variations in scale, orientation and visibility
- Variability within Specificity
- Object of interest has to be recognized in context of multiple other objects and cluttered background

Imaging



A picture worth a thousand words

Video

Sequence of Frame, typically 30 frame per second (fps)

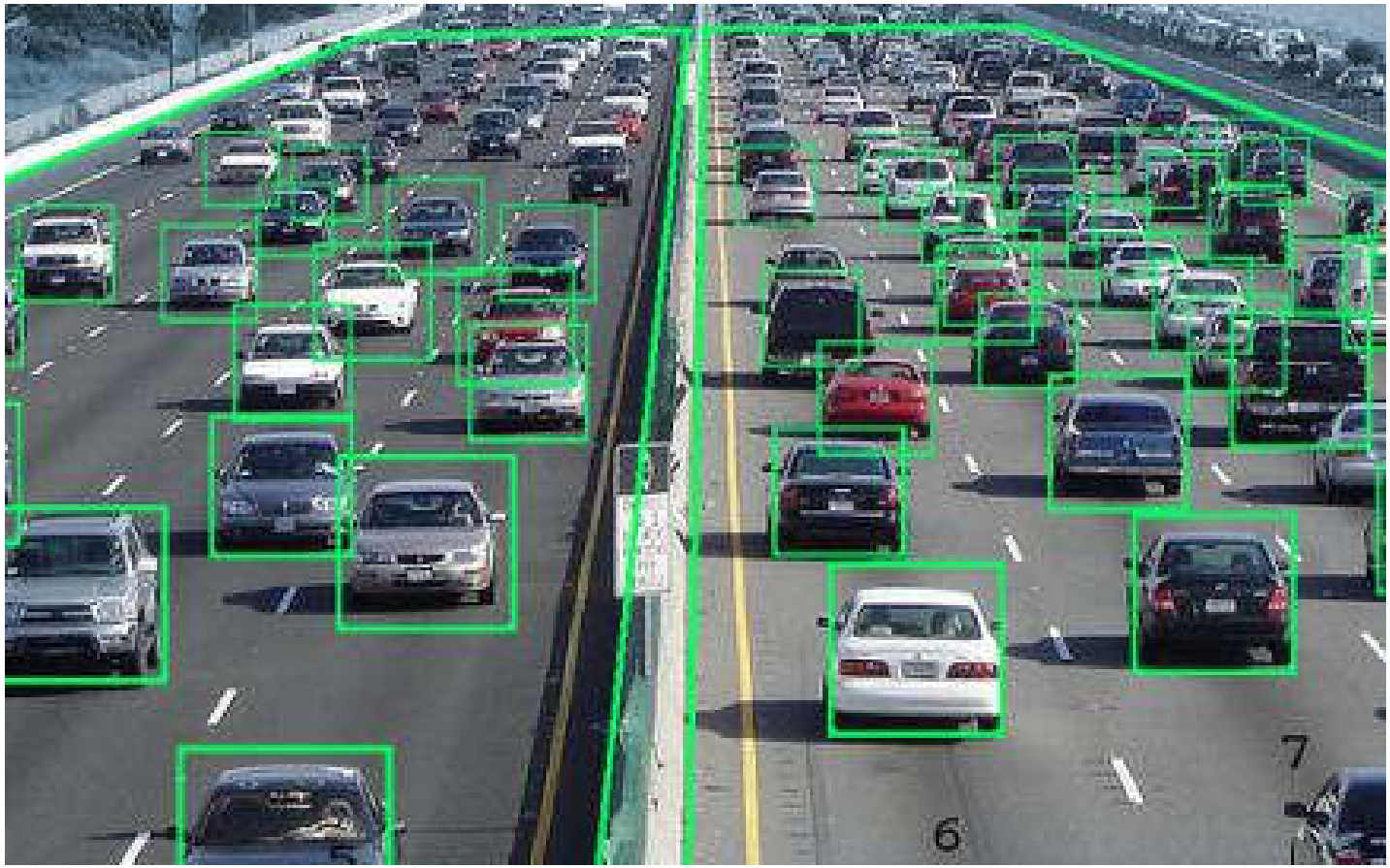
Nowadays


sky
building
STOP
sky
WEST
tree
sky
building
tree
building
tree
person
tree
building
tree
tree
tree
tree
car
STOP
car
car
car
car
car
tree
car
car
car
car
car
car
car
road
car
car
road
car
road
road
building
car
road
road
grass
person

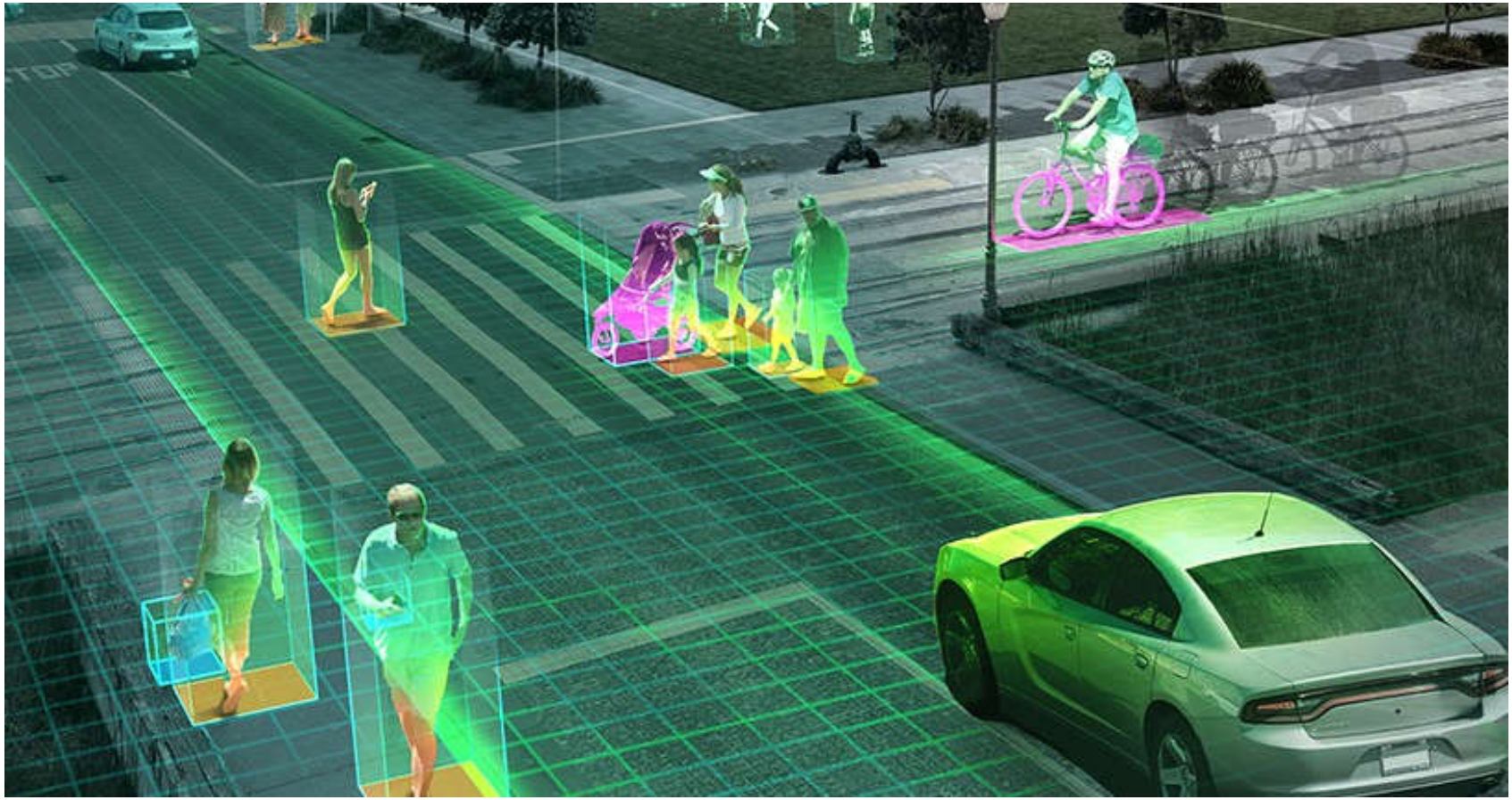
Nowadays


building
sky
STOP
sky
WEST
tree
sky
tree
building
building
building
tree
person
tree
tree
building
tree
tree
tree
tree
car
car
tree
car
car
car
car
car
car
car
car
car
road
car
road
car
car
road
road
building
car
road
road
grass
person

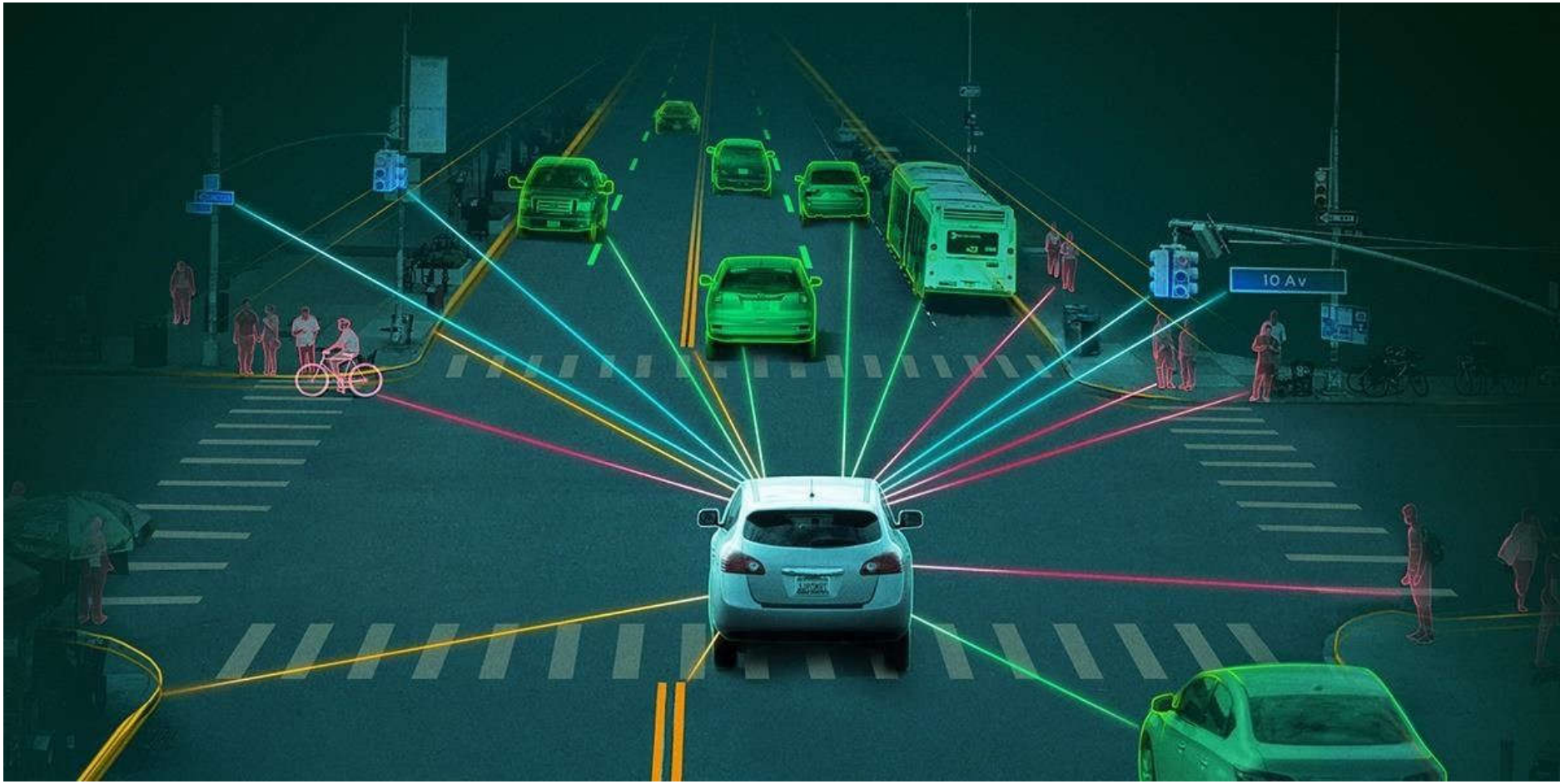
Nowadays



Nowadays



Nowadays



Now days


Stone!
A foreign object has been detected.
Trigger expulsion.

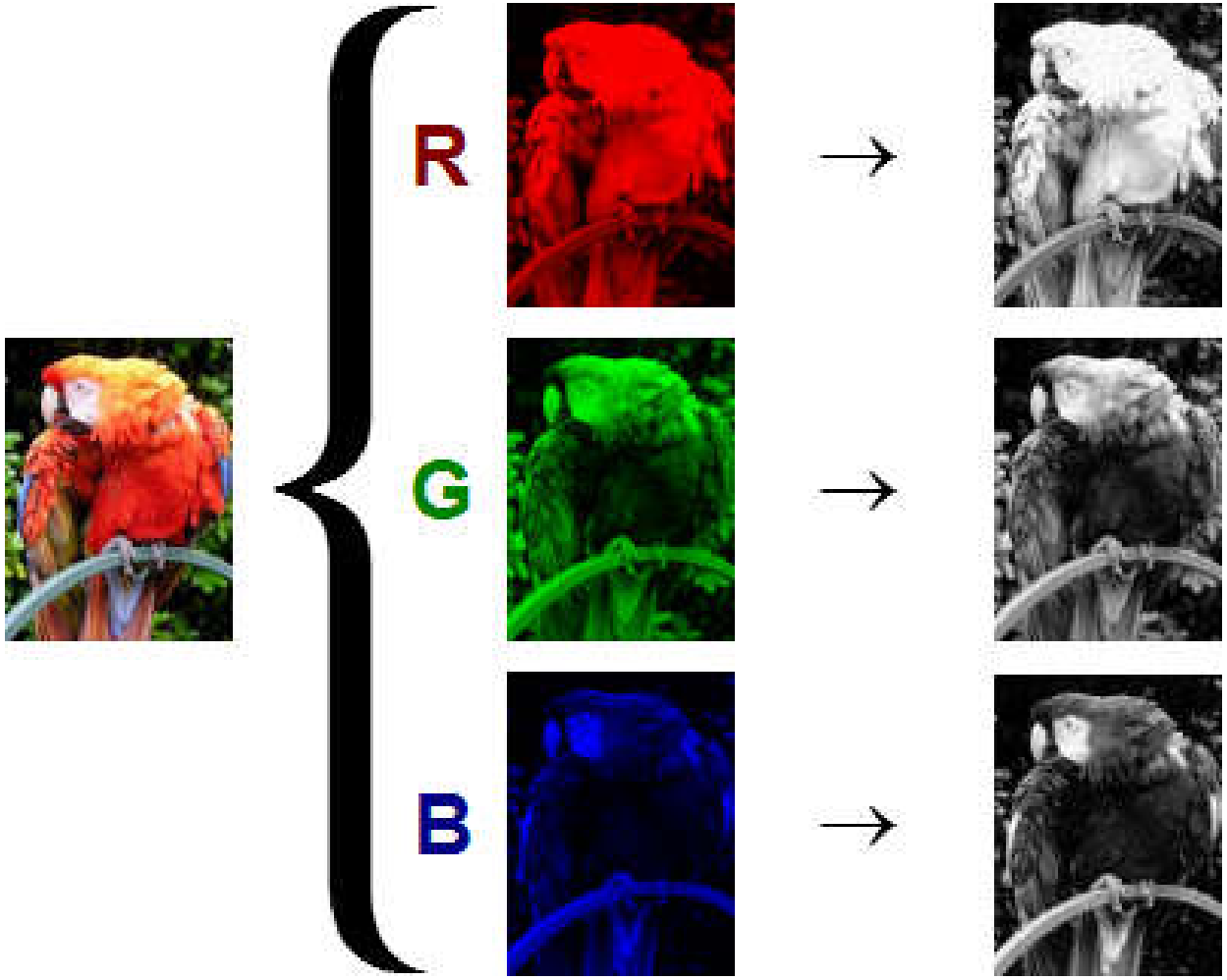
What is vision?

- Recognize objects
  - people we know
  - things we own
- Locate objects in space
  - to pick them up
- Track objects in motion
  - catching a baseball
  - avoiding collisions with cars on the road
- Recognize actions
  - walking, running, pushing

Image

- An image *I* is a 2D array or matrix, where each element in the array is intensity value or gray level)

Gray Image



$$I = \{I_R, I_G, I_B\}$$

Image Resolution

- How many Rows and Column in the image

128 × 128

512 × 512

34	23	58	89	106	97	89	83	83	81
97	39	23	67	75	89	89	89	89	81
139	73	26	67	67	58	75	81	81	75
141	147	94	106	64	7	23	58	81	83
56	89	147	155	114	73	48	58	73	81
23	64	115	148	155	114	48	25	48	73
23	56	74	81	73	64	73	81	89	89
73	56	45	62	57	56	73	81	82	82
97	64	81	106	116	97	89	82	82	82
97	81	89	86	89	97	81	78	82	97

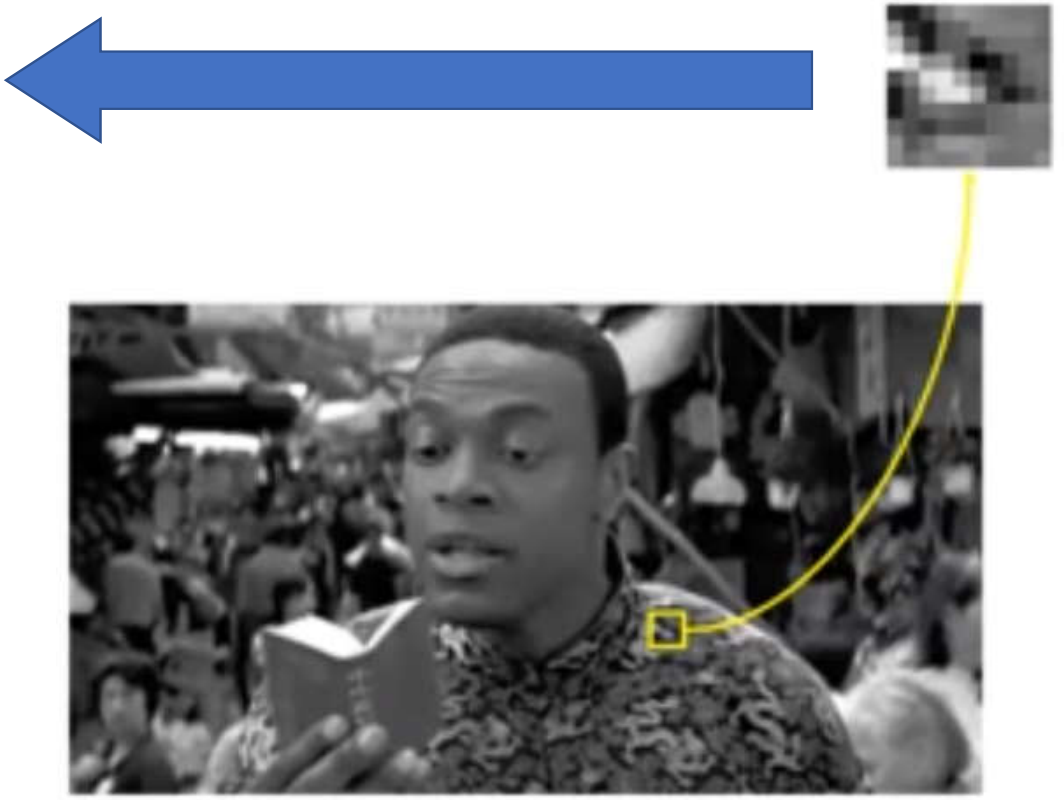


Image Format

- JPEG
- PNG
- BMP
- TIF

Video Format

- AVI
- MPEG

Image Formation

- Light source
- Camera Parameters
- Scene
  - Surface Reflectance
  - Surface Shape

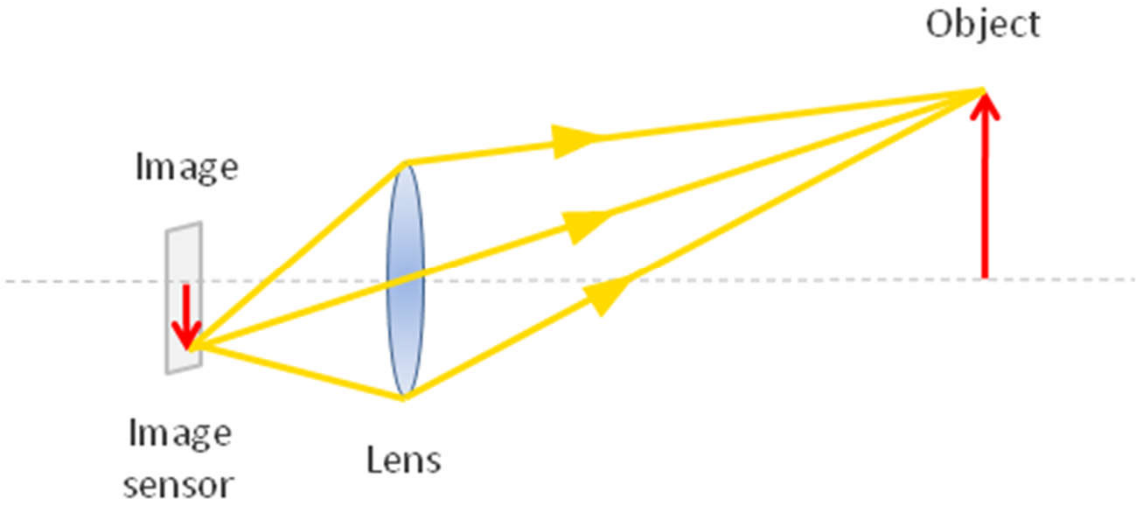


Image Formation

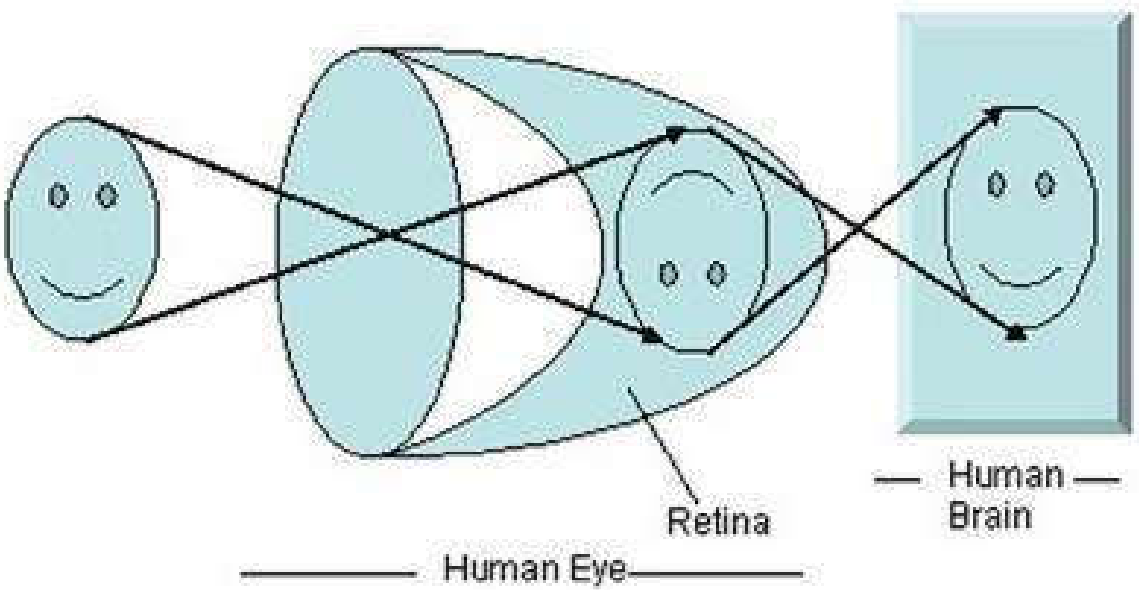
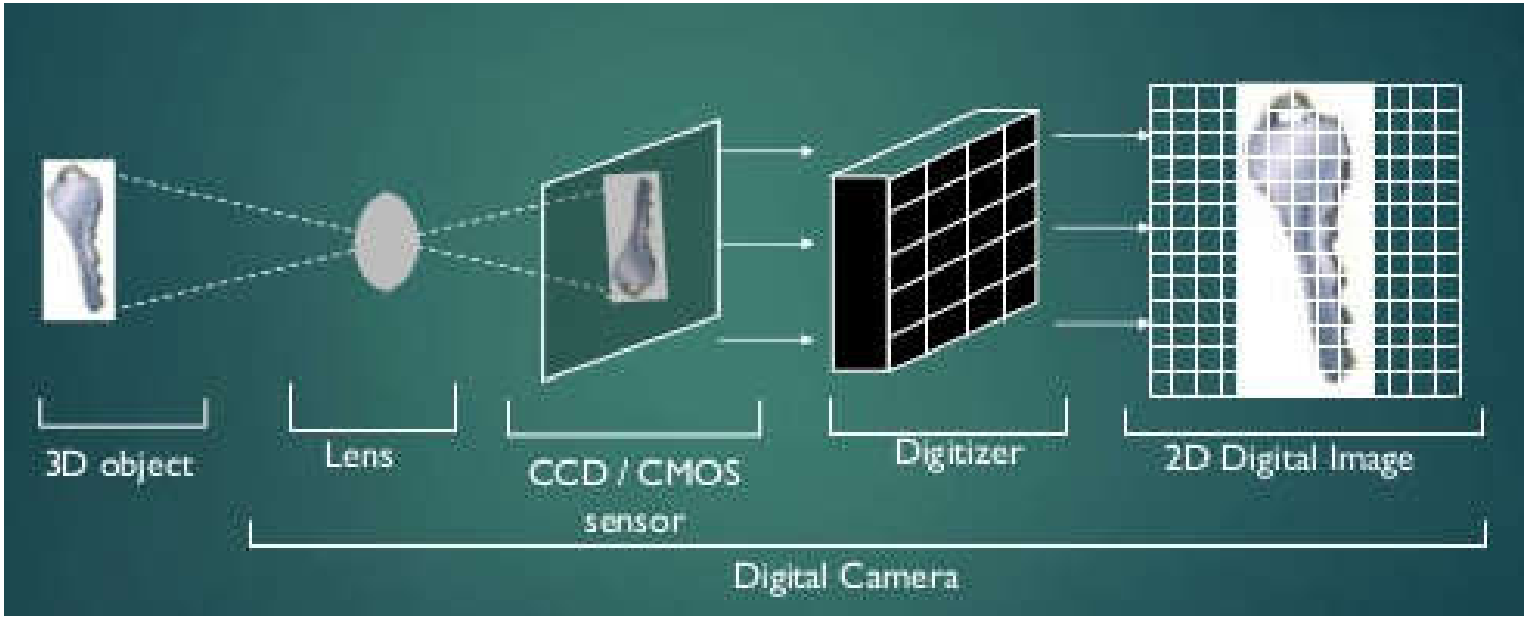
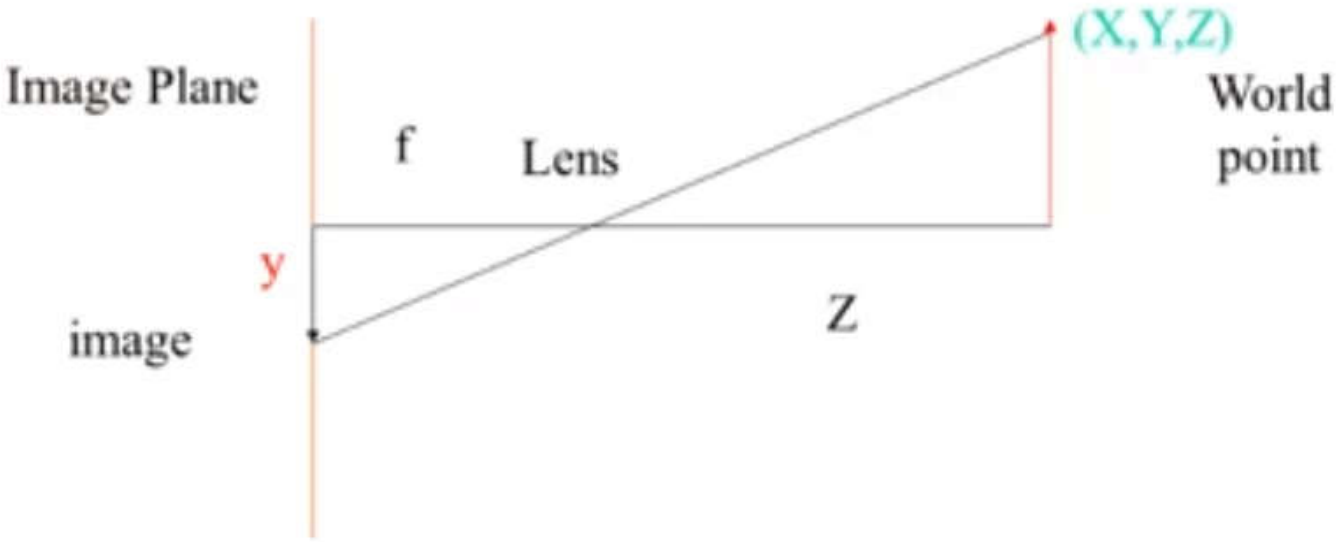


Image Formation





- $-\frac{y}{Y} = \frac{f}{Z} \qquad \Rightarrow \qquad y = -\frac{fY}{Z}$
- $x = -\frac{fX}{Z}$

- Computer vision application trying to recover 3D images from 2D images and it is called 3D shape and it is called ***shape from X***
  - *Stereo*
  - *Motion*
  - *???*


DSC06903_1500x.jpg


DSC06904_1500x.jpg


DSC06905_1500x.jpg


DSC06906_1500x.jpg


DSC06907_1500x.jpg


DSC06908_1500x.jpg


DSC06909_1500x.jpg


DSC06910_1500x.jpg


DSC06911_1500x.jpg


DSC06912_1500x.jpg


DSC06913_1500x.jpg


DSC06914_1500x.jpg


DSC06915_1500x.jpg


DSC06916_1500x.jpg


DSC06917_1500x.jpg

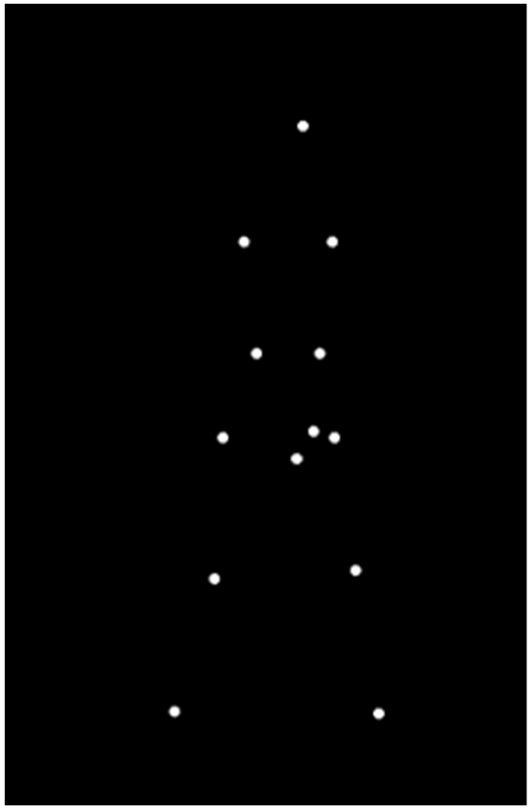

DSC06918_1500x.jpg


DSC06919_1500x.jpg


DSC06920_1500x.jpg


DSC06921_1500x.jpg

Moving Light Display

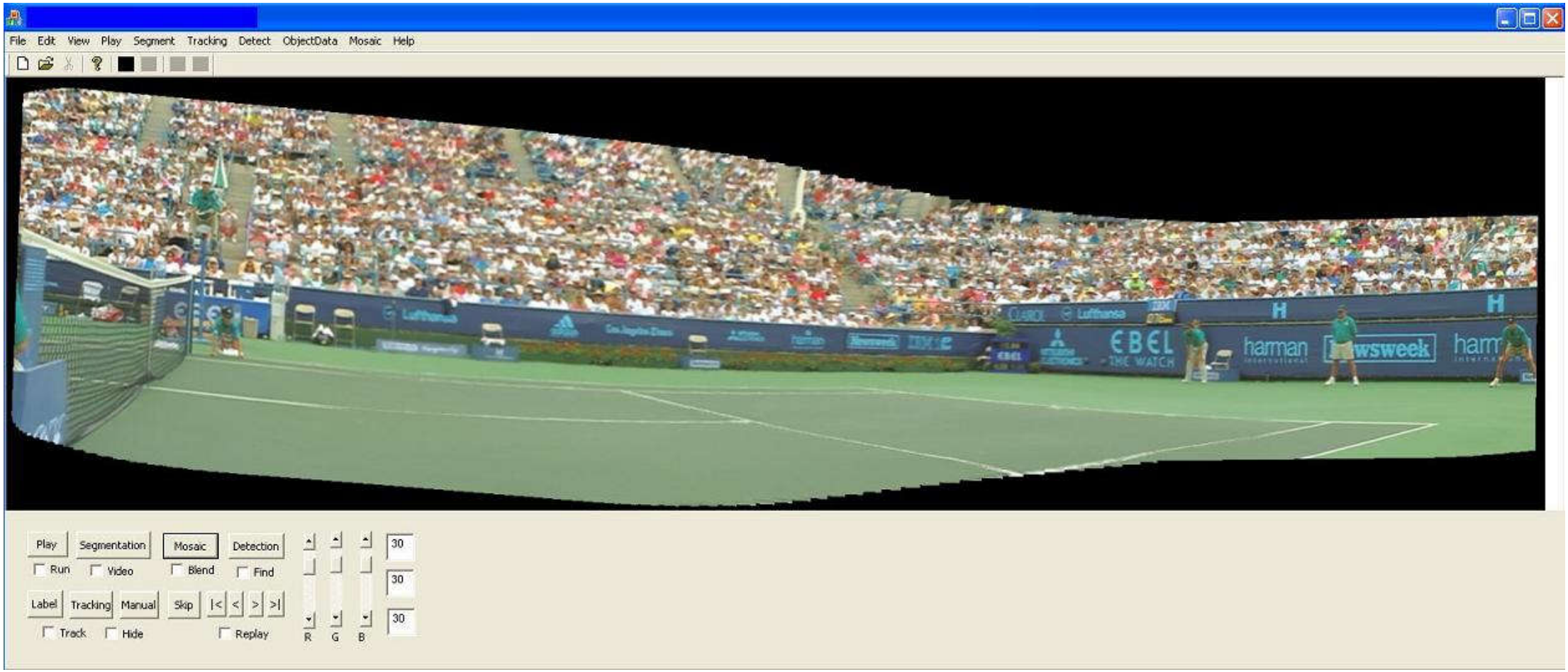


Useful in dark scenes

Flow motion

- **https://www.youtube.com/watch?v=GIUDAZLfYhY**

Video clip ad Mosaic



Application

- Object Detection
- Object Recognition
- Face Recognition
- Video Surveillance and Monitoring
  - Detection and Tracking
- Robotics
- Unmanned Aerial Vehicle (UAV)

Next

- Human Vision, Color Spaces and Transforms

References

- Computer Vision: Algorithms and Applications, Rick Szeliski, 2010.
- Fundamentals of Computer Vision, Mubarak Shah, 1997.
- Feature extraction image processing for computer vision, Nixon, Mark S and Aguado, Alberto S, 2012, Academic Press.
- Color image processing: methods and applications, Lukac, Rastislav and Plataniotis, Konstantinos N, 2006, CRC press.