Computer Vision

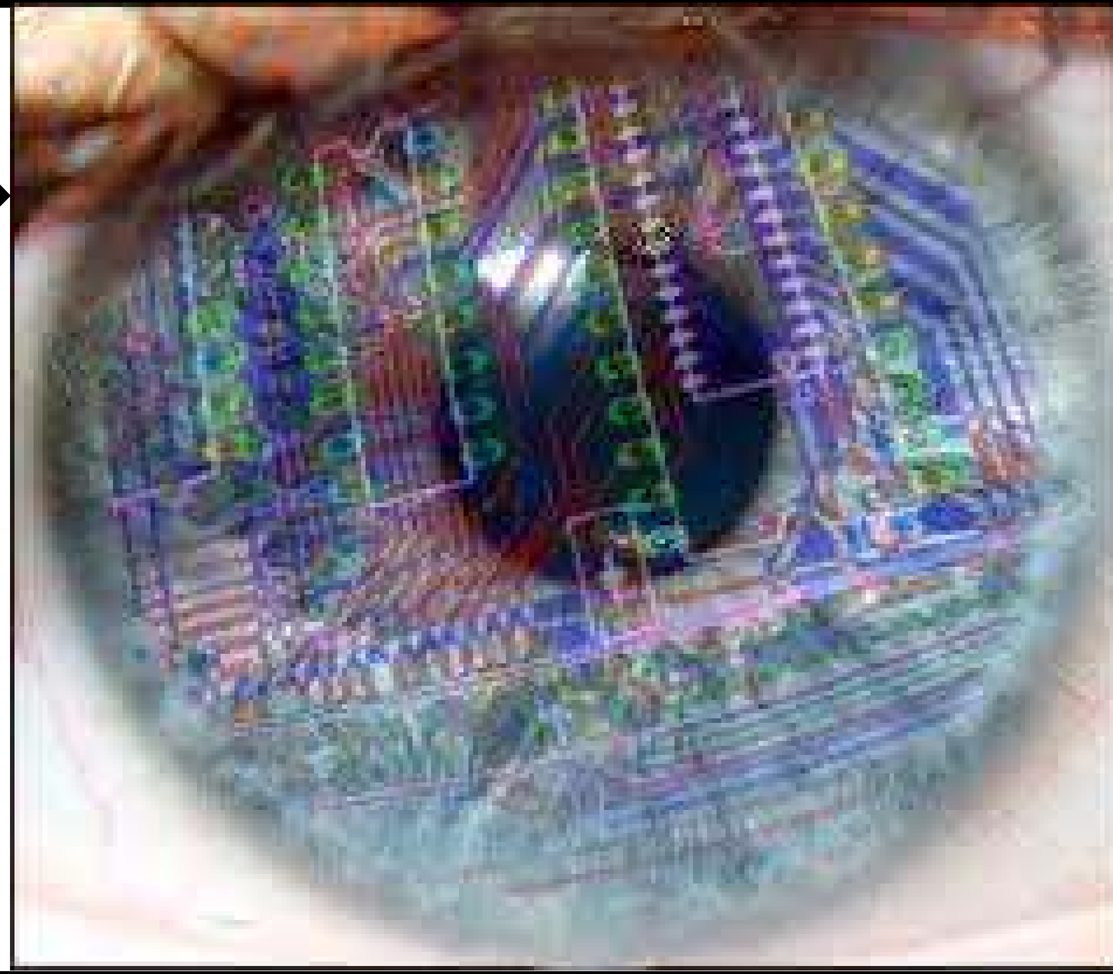
Dr Ankur Jain

Phd (CPS), MTech (CSE), MBA (IT)

Assistant Professor, (SCSE)

VIT Bhopal University, Bhopal-Indore Highway, Kothrikalan, Sehore , Madhya Pradesh - 466114
Email: ankurjainjob@gmail.com, ankur.jain@vitbhopal.ac.in

Mobile: 7415259169



My Teaching Style

- Application based study goal
- Objective based lecture presentation
- Example based topic explanation
- Step by step either implementation or simulation or emulation
- Top down approach for broad topic introduction and review
- Bottom up approach for design, simulation and implementation

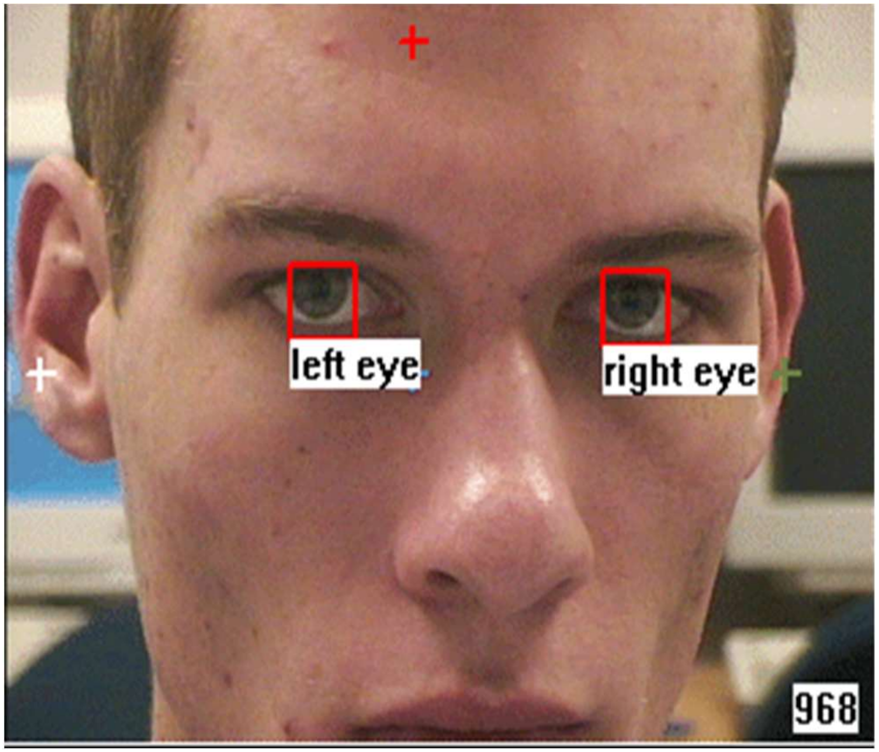
Subjects of interest

- Cyber physical system
- Internet of things
- Industrial data communication
- Data acquisition, logging and analysis
- Embedded systems
- Industrial automation and Industry 4.0
- PLC, DCS and SCADA systems
- Machine learning
- Deep learning
- Reinforcement learning
- Intelligent systems
- Computer networks
- Programming and OOPS
- Control system design
- Computer vision
- Image Processing
- Analog electronics
- Digital electronics
- Power electronics
- Microcontroller and microprocessor
- Robotics
- Optimization
- Raspberry pi, Arduino, etc.

Note: I am not well versed in all the topics. I am a student learning day by day according to the need of the project.

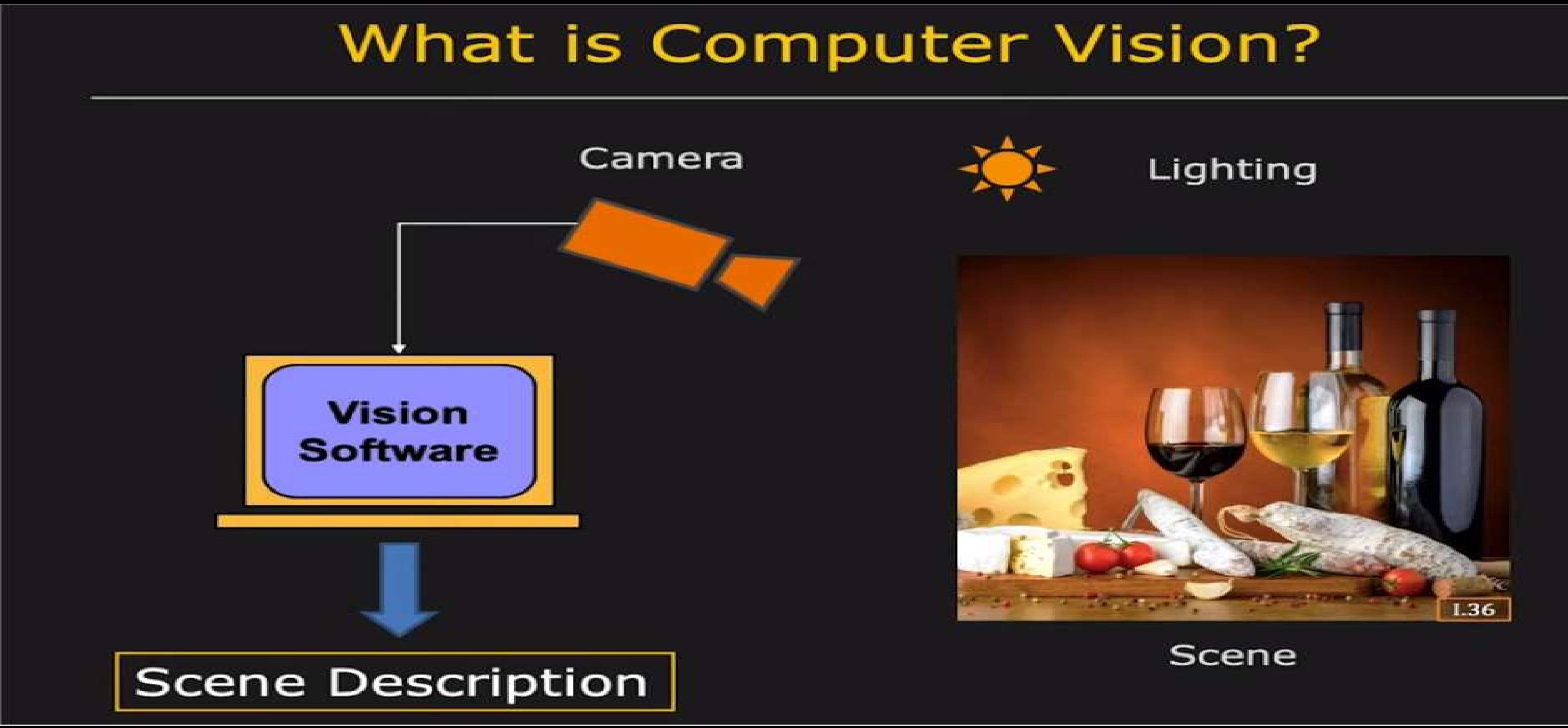
Marks Distribution

1. End Term : 40%
2. Mid Term : 30%
3. Project : 10%
4. Project Report: 5%
5. Paper Presentation: 5%
6. Quiz (Class Test): 5%
7. Attendance: 5%



Tutorial objectives

- Learn by Doing
- Introduction
- Marks Distribution
- Group Formation
- Project Selection
- Installation of tools

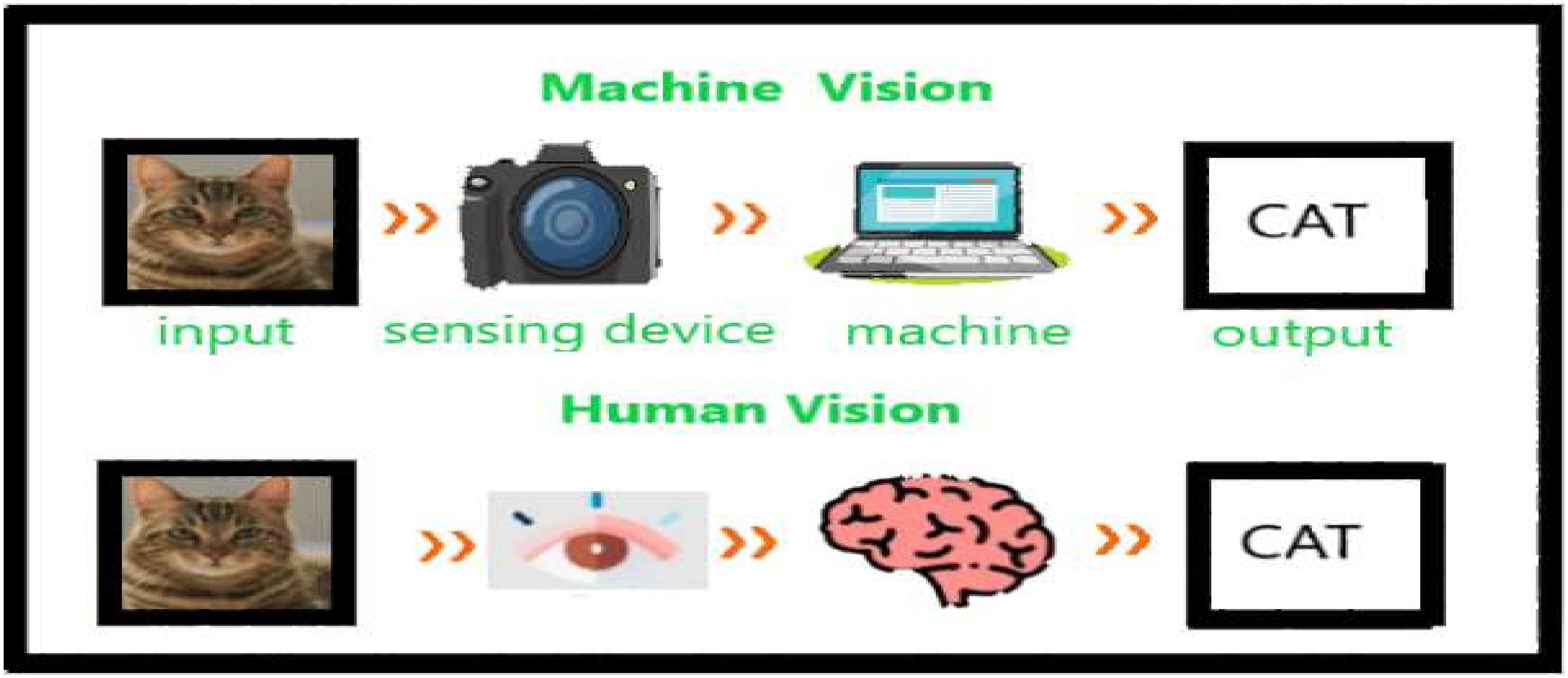

What is Computer Vision?
Camera
Lighting
Vision Software
Scene Description
Scene
1.36

Introduction : What is computer vision

- **Computer Vision "is a discipline that studies how to reconstruct, interpret and understand a 3D scene from it's 2D images in terms of the properties of the structures present in the scene".**

- **Computer vision systems analyze images and video automatically and determine what the computer "sees" or "recognizes."**

Machine vision vs Human Vision



Difference between computer vision

Computer Vision	Human Vision
Computer vision allows computer to sense their surroundings and identify things, similar to how human vision perceive things.	Humans perceive the things as they are and retain what they recognize, storing it deep in the brain until they come across those things again.
Computer vision uses machine learning techniques and algorithms to identify, distinguish and classify objects.	Human vision is about how eyes detect light patterns and coordinate with the brain to translate light into images.
Object recognition is one of the most challenging problems in computer vision.	Humans recognize objects effortlessly, and have no problems describing objects in a scene.

Continued…

Human Vision	Computer Vision
Multitasking	Task-oriented
Visible patterns	Invisible patterns
Supervision tasks	Periodic tasks
Computer interaction	Human interaction

Computer Vision and Nearby Fields

- Computer Graphics: Models to Images
- Comp. Photography: Images to Images
- Computer Vision: Images to Models

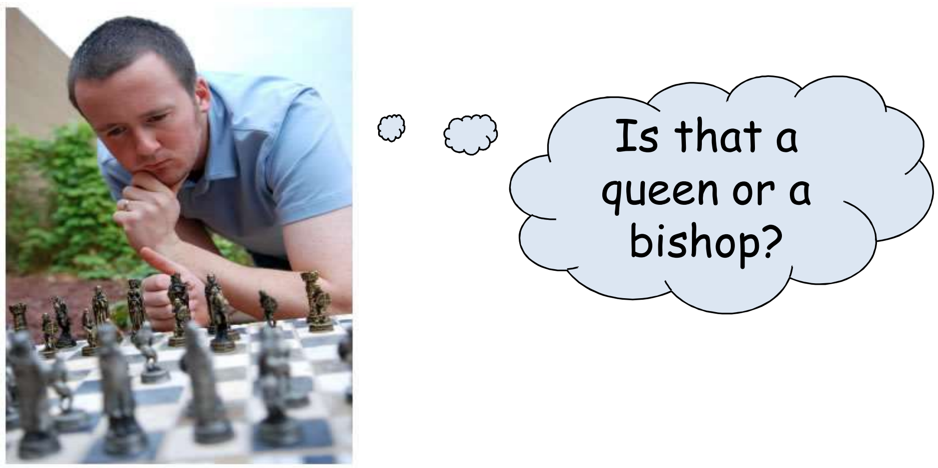
Computer Vision

Make computers understand images and video.

- What kind of scene?
- Where are the cars?
- How far is the building?
- …

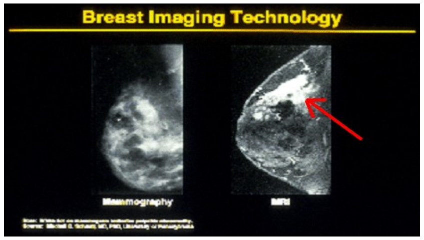
Vision is really hard

- Vision is an amazing feat of natural intelligence
  - Visual cortex occupies about 50% of Macaque brain
  - More human brain devoted to vision than anything else



Why computer vision matters

Safety

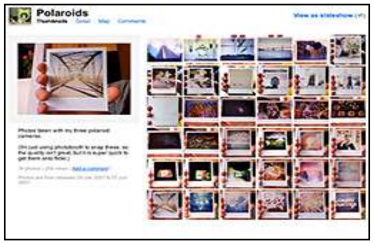


Health

Security

Comfort

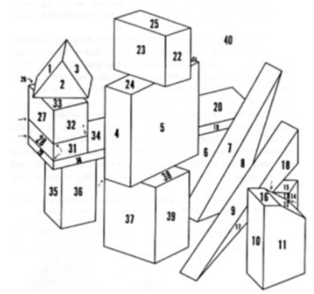
Fun



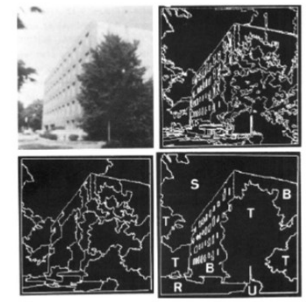
Access

Ridiculously brief history of computer vision

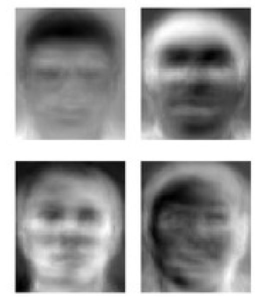
- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960’s: interpretation of synthetic worlds
- 1970’s: some progress on interpreting selected images
- 1980’s: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990’s: face recognition; statistical analysis in vogue
- 2000’s: broader recognition; large annotated datasets available; video processing starts



Guzman ‘68



Ohta Kanade ‘78

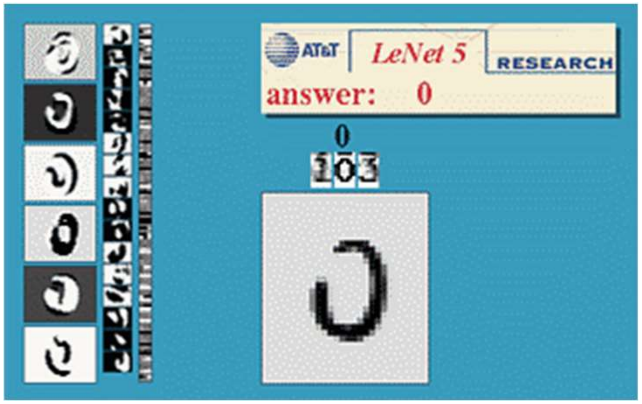


Turk and Pentland ‘91

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



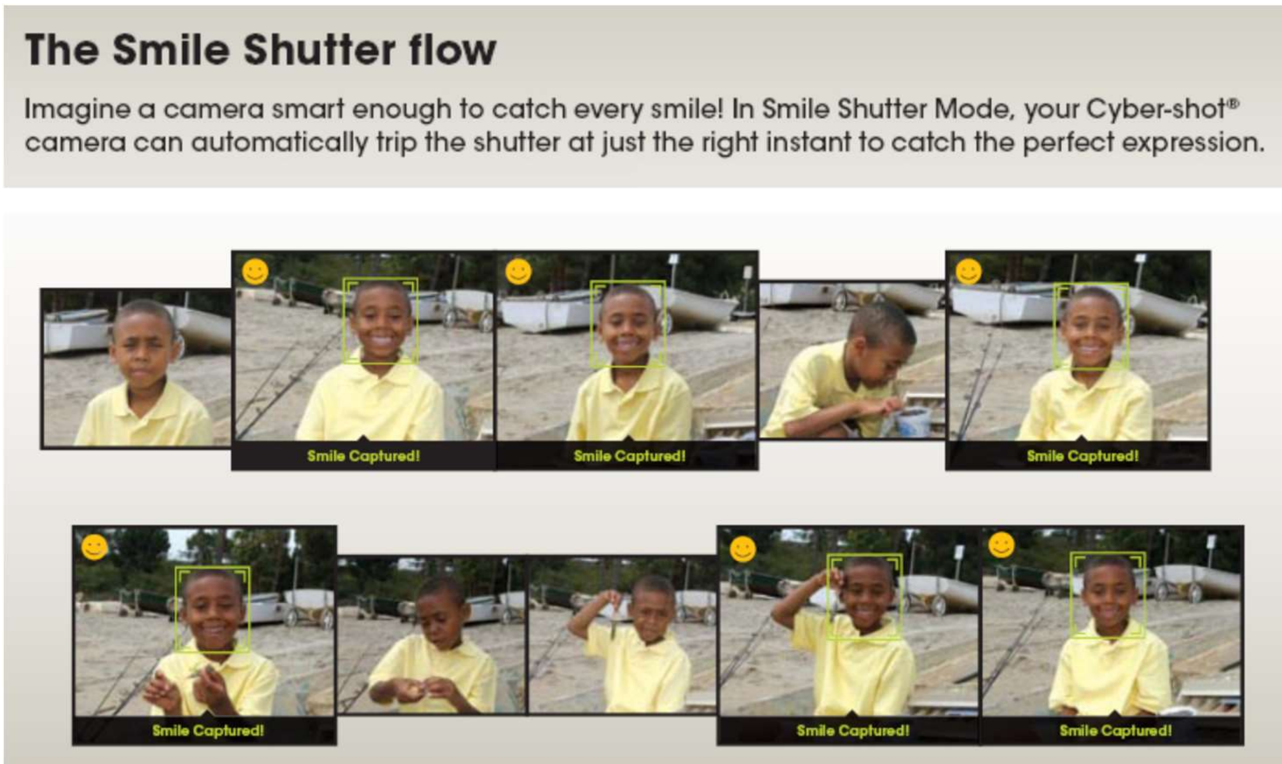
Digit recognition, AT&T labs
http://www.research.att.com/~yann/

License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition

Face detection

- Many new digital cameras now detect faces
  - Canon, Sony, Fuji, …

Smile detection



[Sony Cyber-shot® T70 Digital Still Camera](#)

3D from thousands of images

Building Rome in a Day: Agarwal et al. 2009

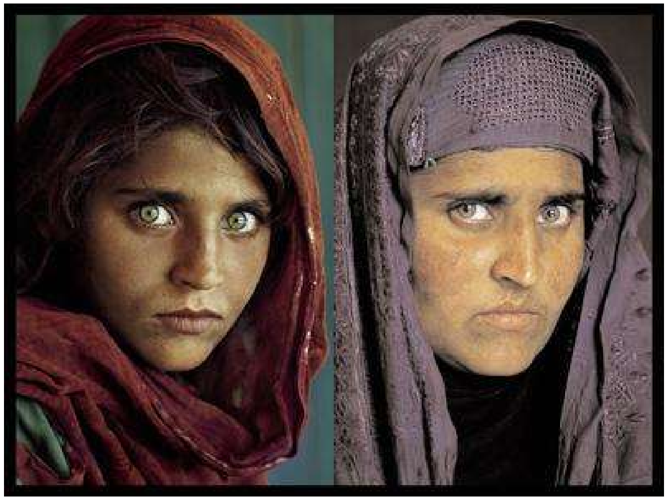
Object recognition (in supermarkets)

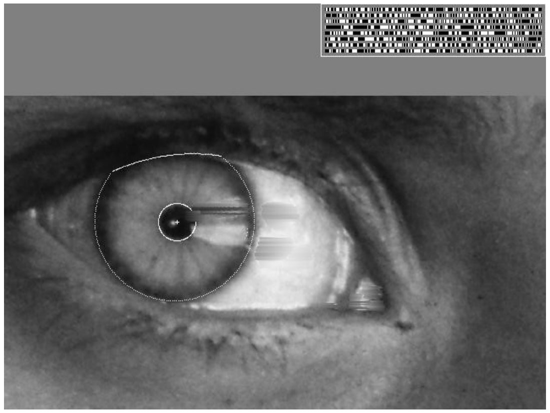
[LaneHawk by EvolutionRobotics]

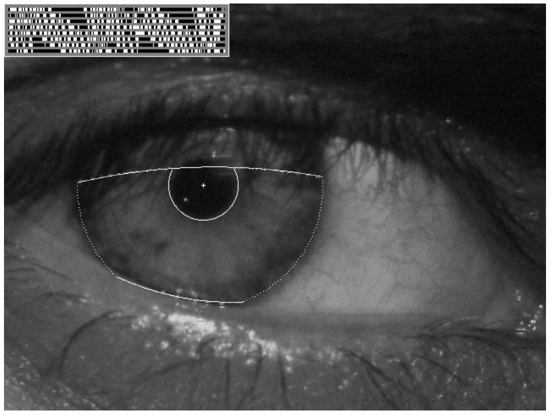
"A smart camera is flush-mounted in the checkout lane, continuously watching for items. When an item is detected and recognized, the cashier verifies the quantity of items that were found under the basket, and continues to close the transaction. The item can remain under the basket, and with LaneHawk,you are assured to get paid for it… "

Vision-based biometrics

"*How the Afghan Girl was Identified by Her Iris Patterns*" Read the story
wikipedia







Login without a password…

Fingerprint scanners on many new laptops, other devices

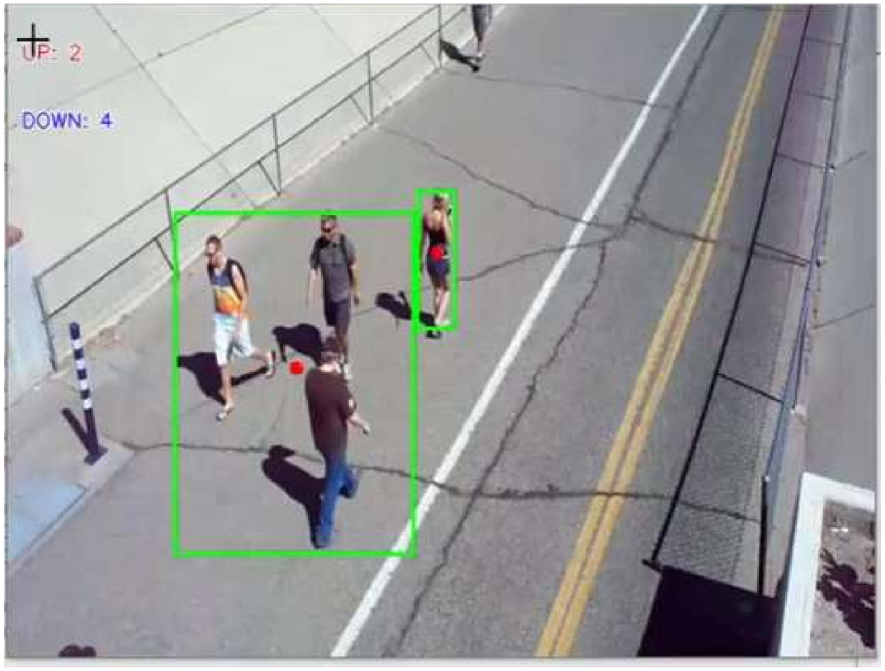
Face recognition systems now beginning to appear more widely
http://www.sensiblevision.com/

Object recognition (in mobile phones)

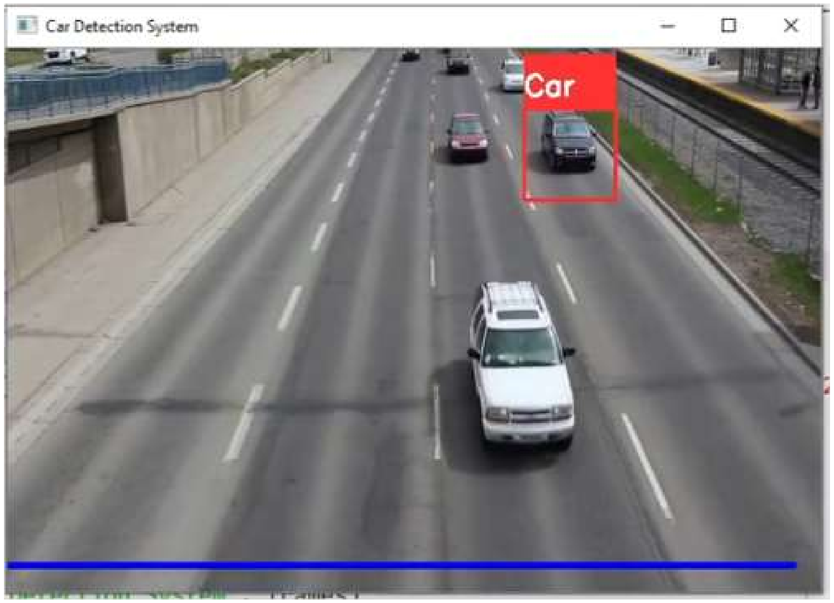
[Point & Find](), Nokia, [Google Goggles,]()

Gesture recognition: Counting of finger

Pedestrian counting



Real Time Vehicle Detection



Special effects: shape capture

The Matrix movies, ESC Entertainment, XYZRGB, NRC

Face Filter

Special effects: motion capture

Pirates of the Carribean, Industrial Light and Magic

Sports

Sportvision first down line
Nice explanation on www.howstuffworks.com

http://www.sportvision.com/video.html

Smart cars

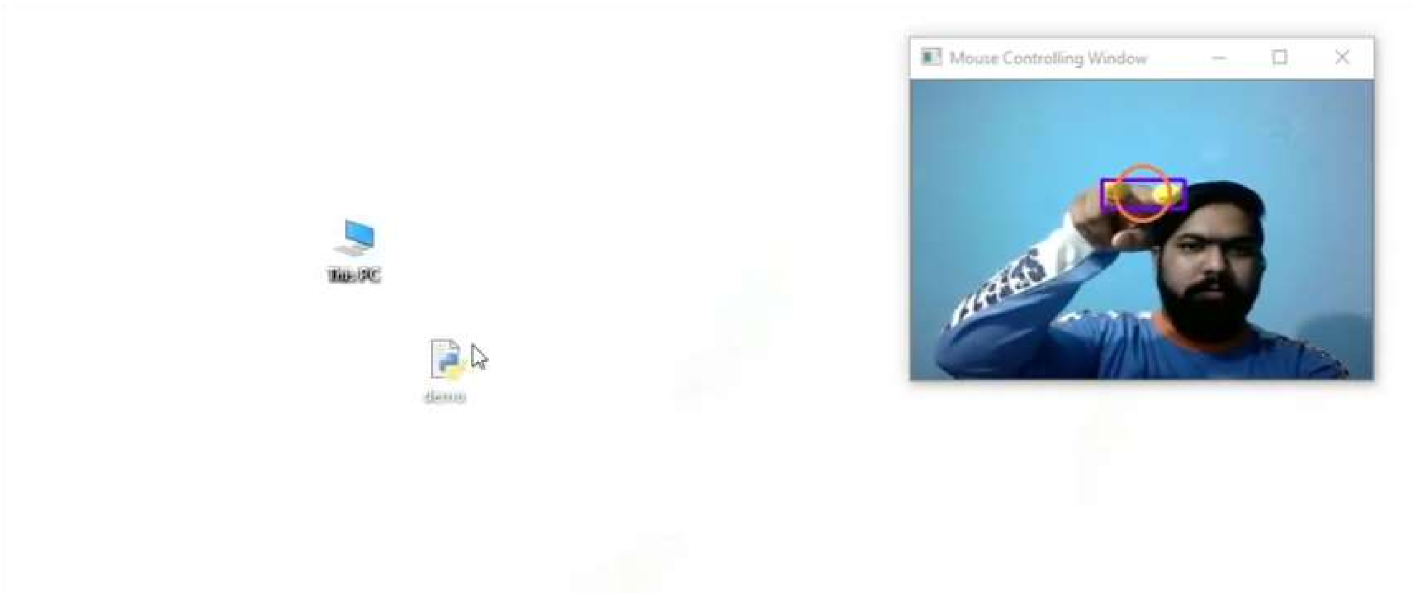
Slide content courtesy of Amnon Shashua

- [Mobileye]
  - Vision systems currently in high-end BMW, GM, Volvo models
  - By 2010: 70% of car manufacturers.

Google cars

http://www.nytimes.com/2010/10/10/science/10google.html?ref=artificialintelligence

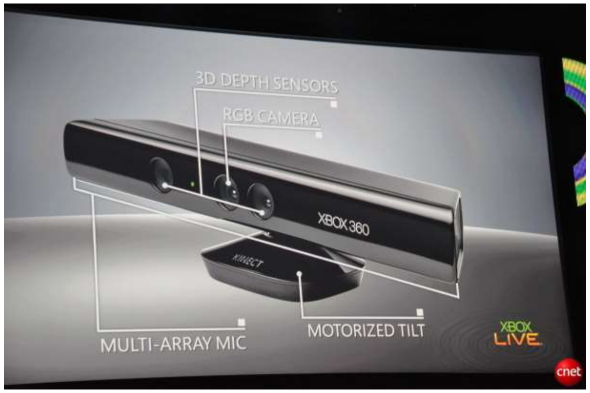
Mouse Control using fingers



Facial Expression Detection

Interactive Games: Kinect

- Object Recognition: http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o
- Mario: http://www.youtube.com/watch?v=8CTJL5lUjHg
- 3D: http://www.youtube.com/watch?v=7QrnwoO1-8A
- Robot: http://www.youtube.com/watch?v=w8BmgtMKFbY



Vision in space

NASA'S Mars Exploration Rover Spirit captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

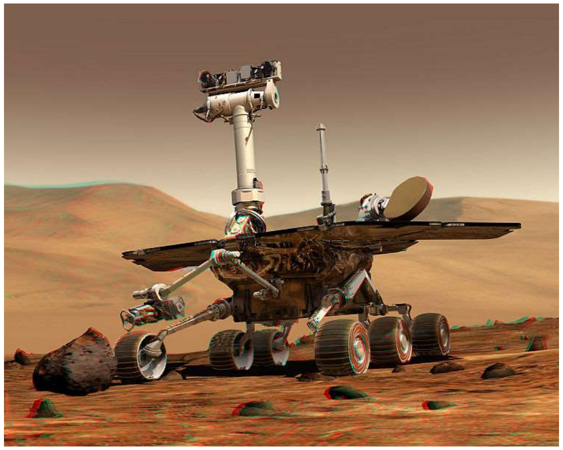
Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read "Computer Vision on Mars" by Matthies et al.

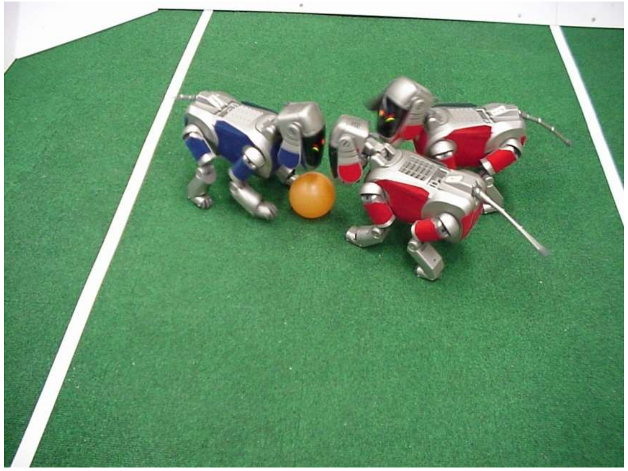
Industrial robots

Vision-guided robots position nut runners on wheels

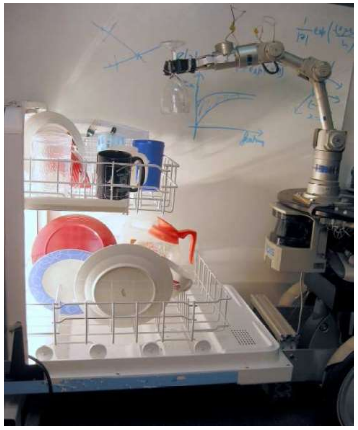
Mobile robots



NASA's Mars Spirit Rover
http://en.wikipedia.org/wiki/Spirit_rover

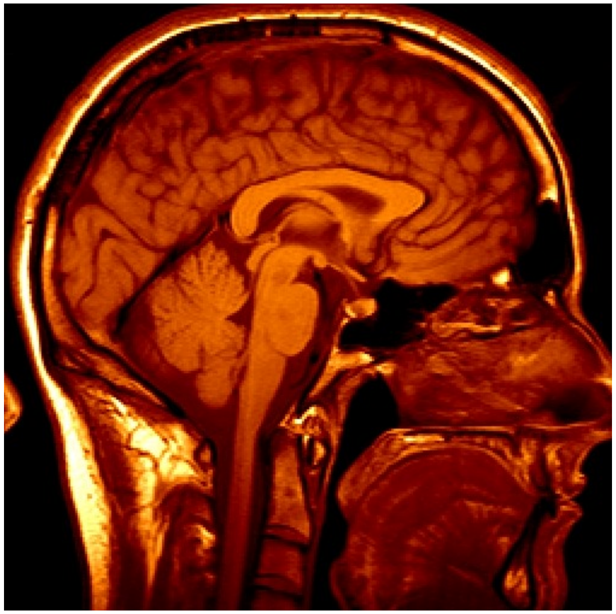


http://www.robocup.org/



Saxena et al. 2008
STAIR at Stanford

Medical imaging



3D imaging
MRI, CT

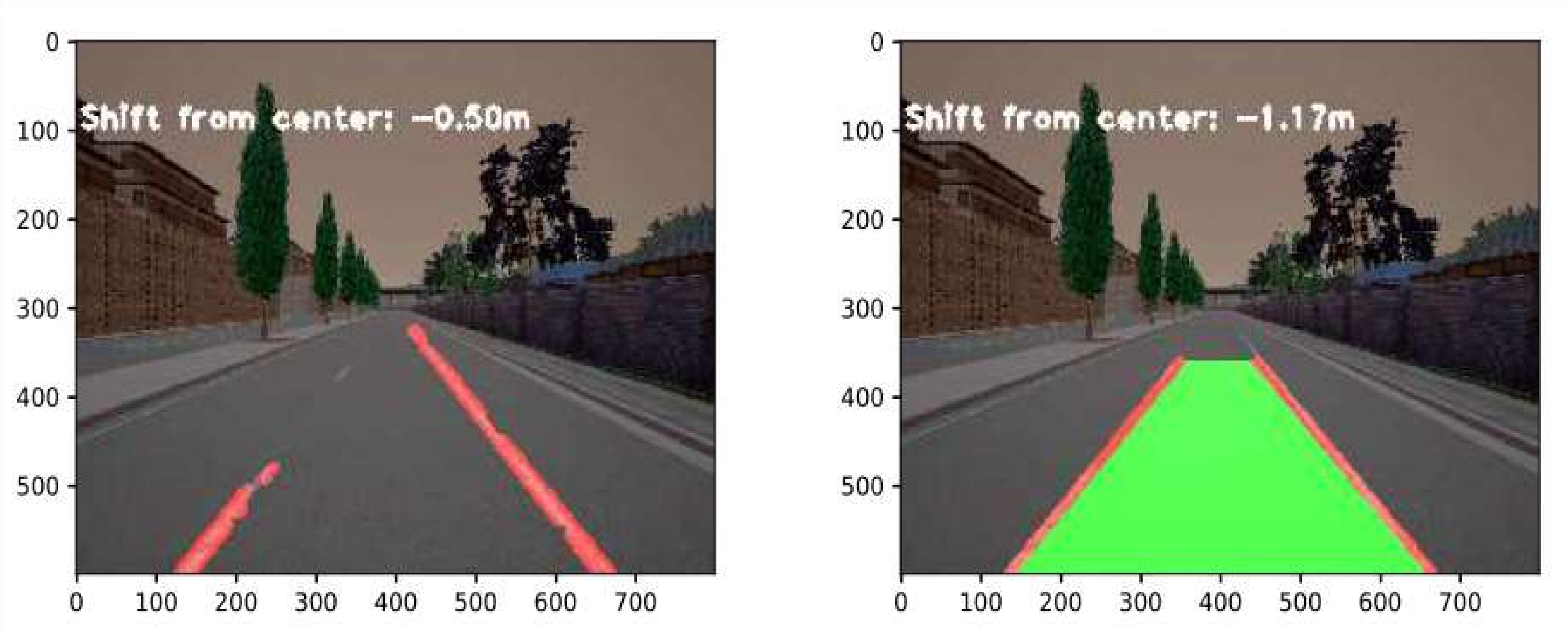
Image guided surgery
Grimson et al., MIT

Face Recognition

Current Research

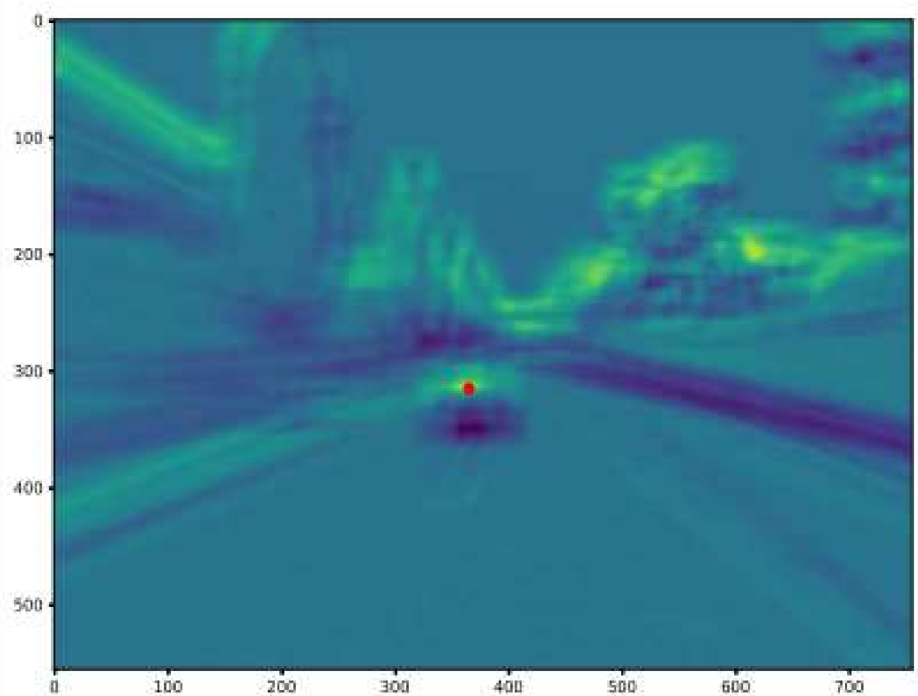
Machine learning: A Computer vision application

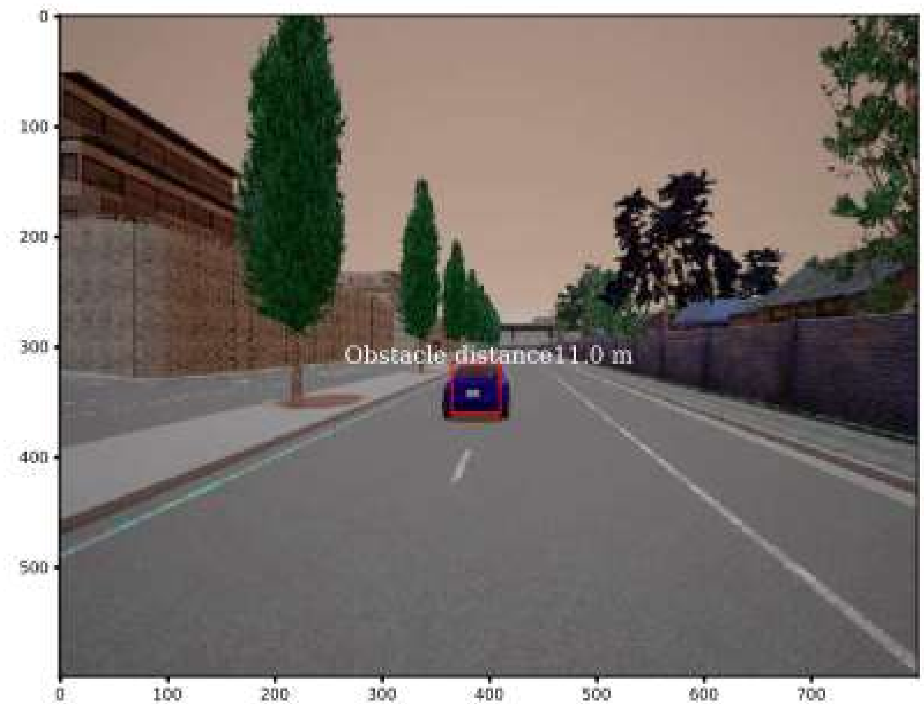
Online lane detection and departure warning system



Object detection, classification and warning alter

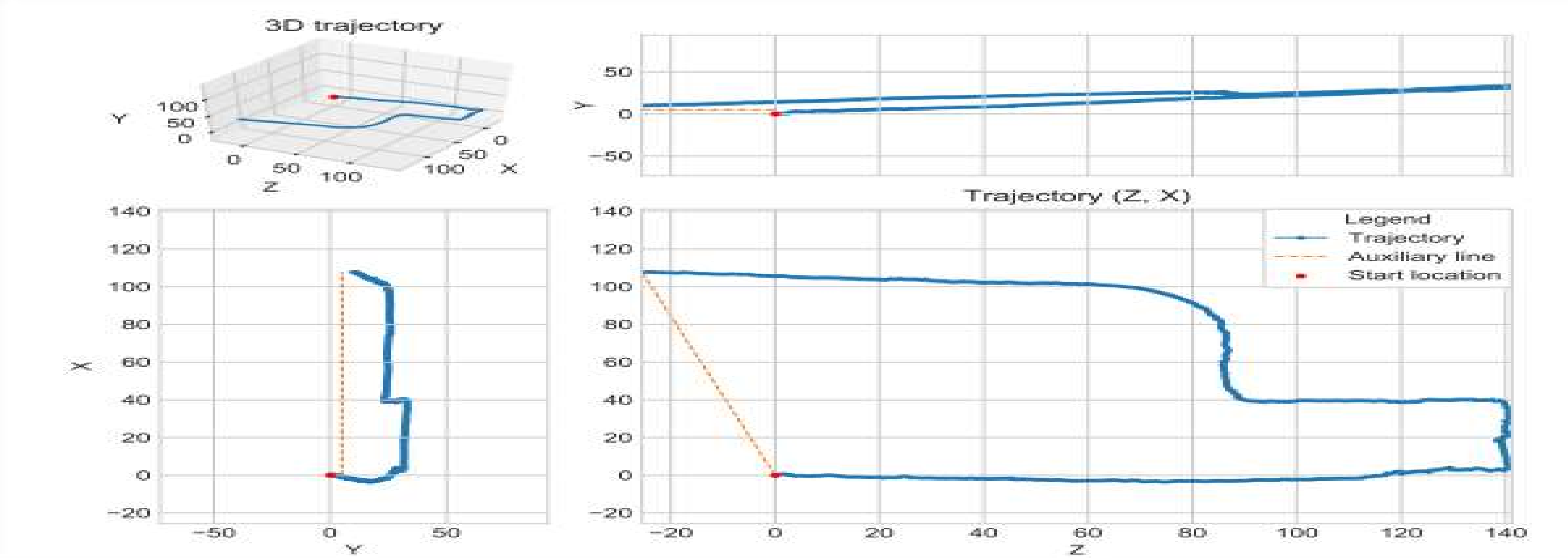
Distance estimation from ego vehicle



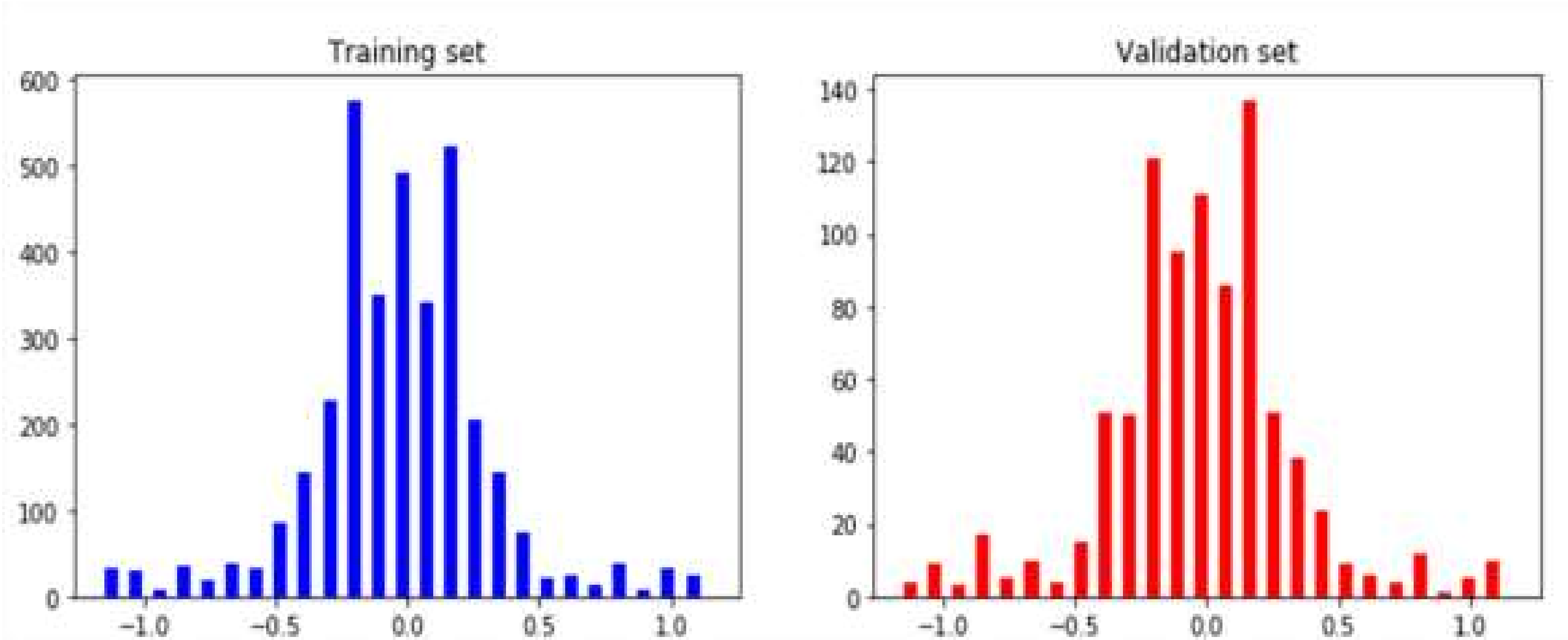


Feature detection and matching

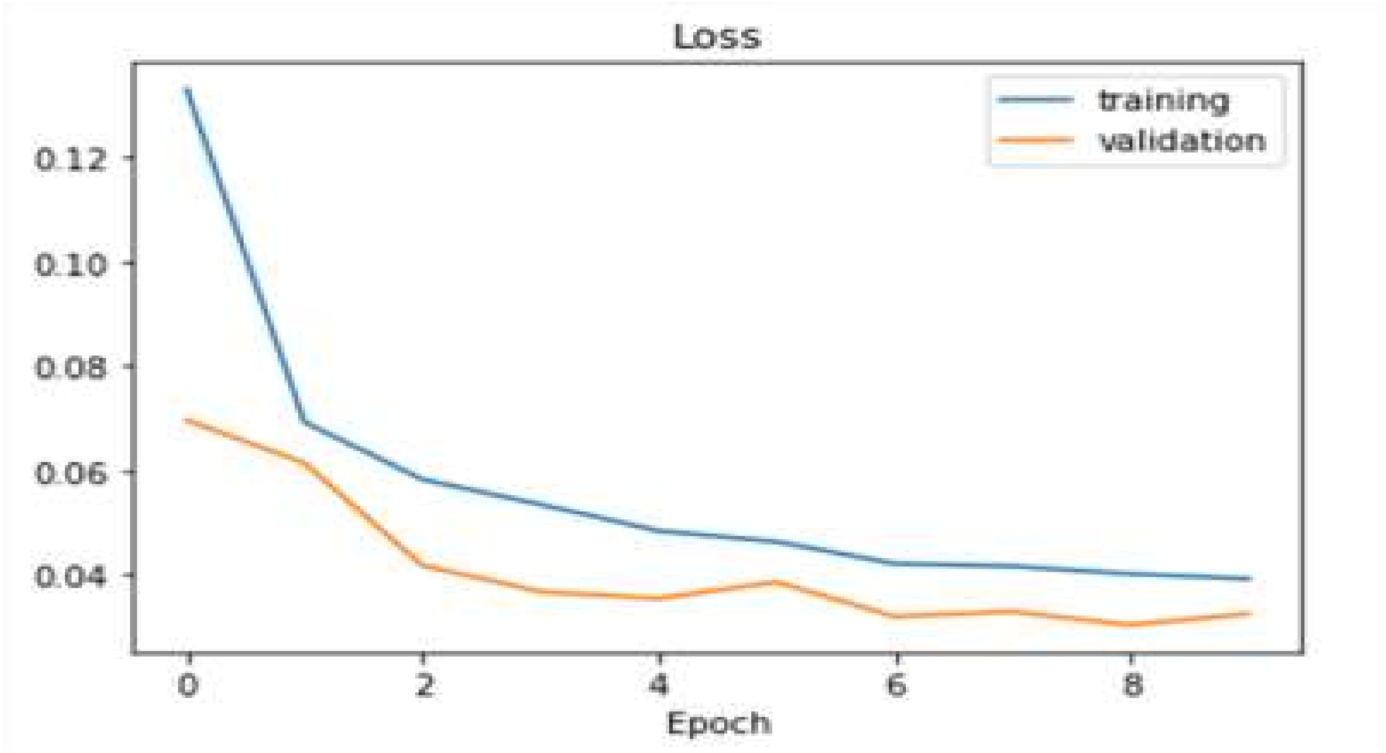
Visual Odometry



Autonomous Driving: Convolution Neural Network



Model efficiency

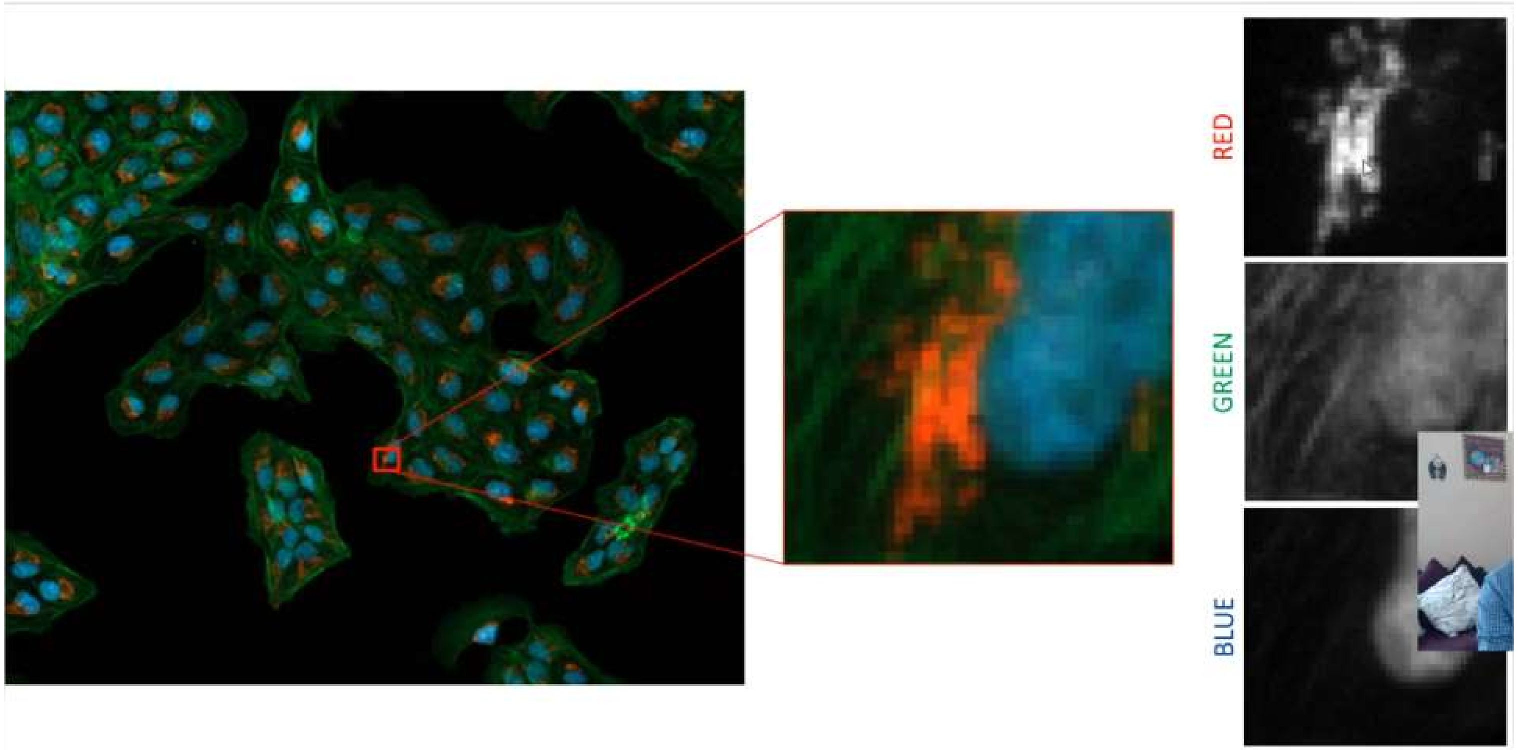


Fundamentals

Introduction

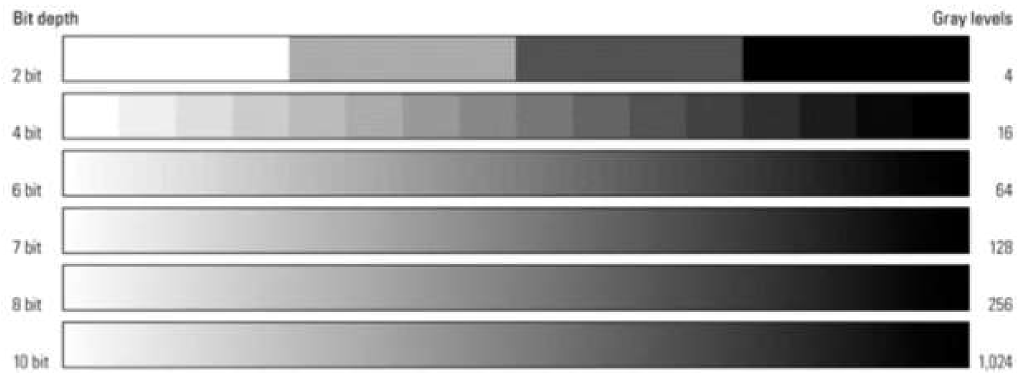
- Image processing
- Computer vision
- OpenCV

What is Digital Image Processing



A pixel in a digital image has a specific value based on the bit depth

Bit depth	Range
1	0 & 1
2	0 to 3
8	0 to 255



Digital image processing involves manipulation of pixel values using an algorithm

Stride = 1

170	245	0	74	149
234	42	64	138	160
32	53	128	202	224
96	117	192	213	21
106	181	255	10	85

Kernel

-1	0	1
2	1	2
1	-2	0

(-1*170) + (0*245) + (1*0) +
(2*234) + (1*42) + (2*64) +
(1*32) + (-2*53) + (0*128) = 394

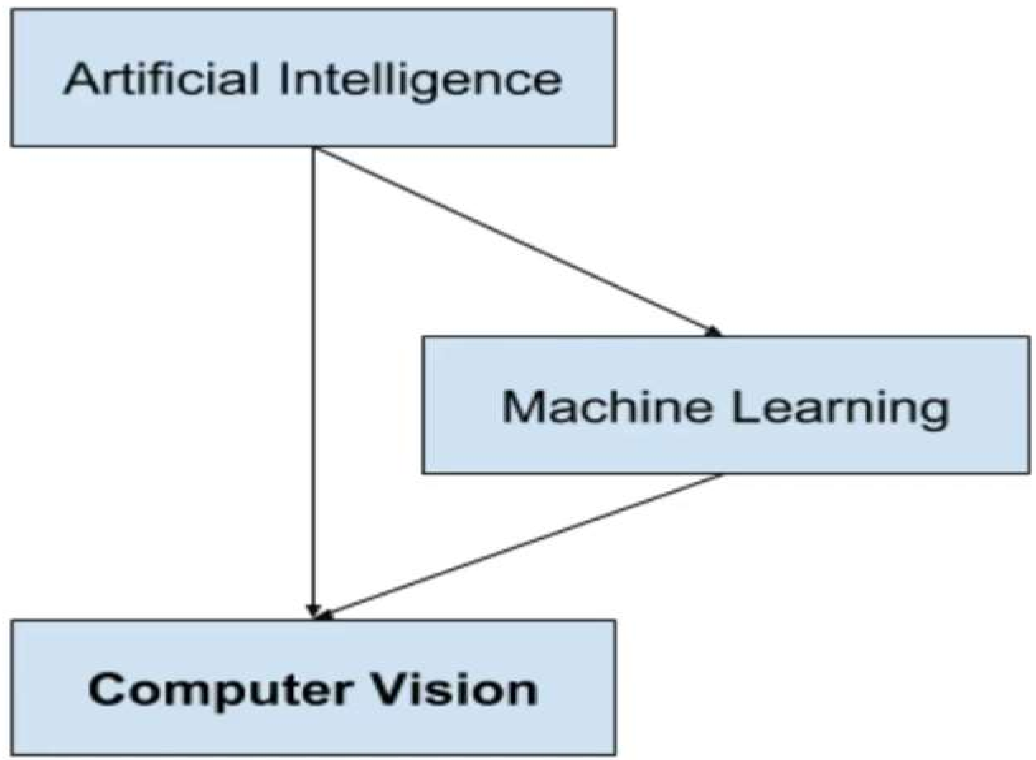
	394			

SINGLE CHANNEL

3 - CHANNEL

What is Video

- Video is the collection of images.
- Image processing is the collection of methods to enhance the quality of image , modify the image or extract some information from the image.
- Computer vision is an interdisciplinary branch that deals with how computer can gain high level understanding of from digital images/video.



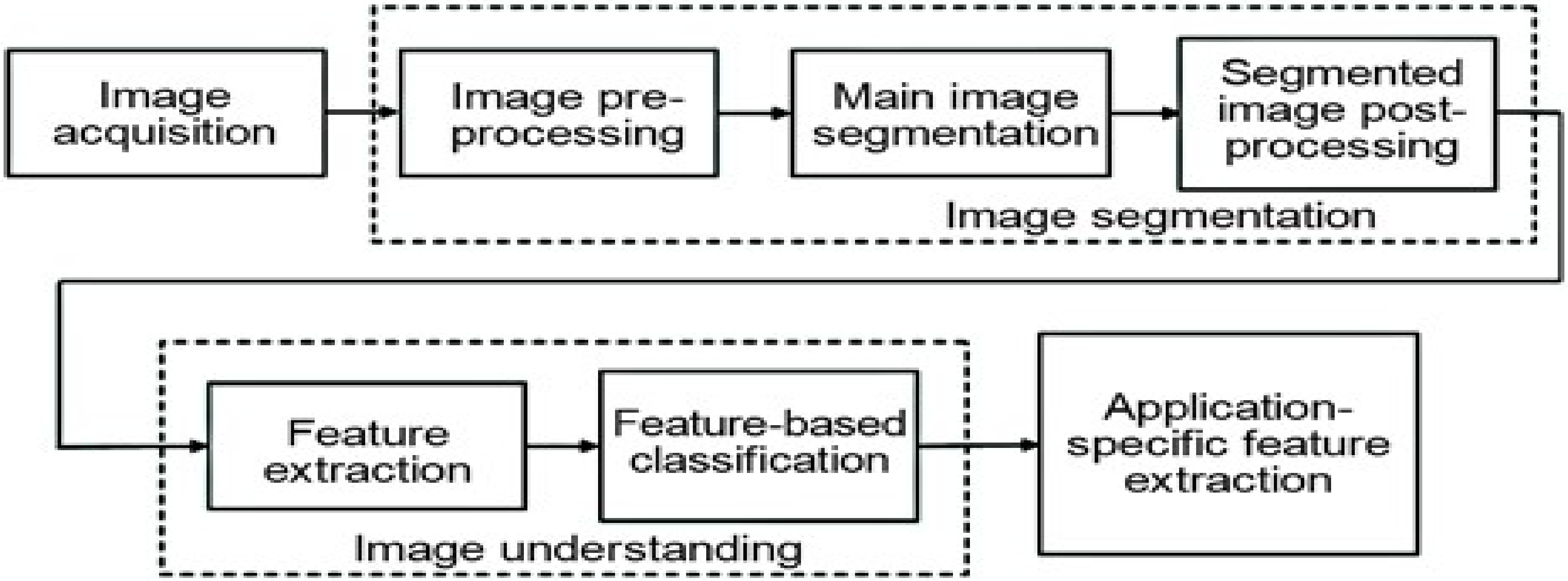
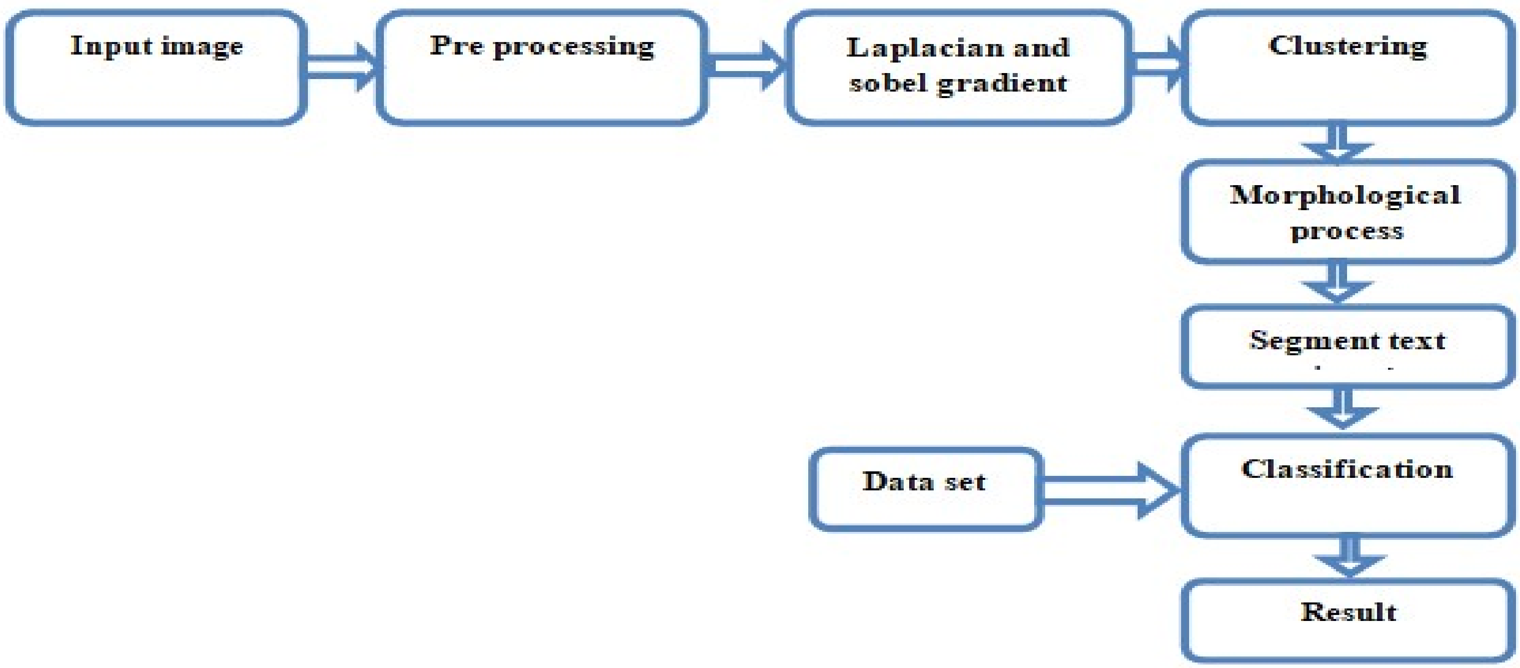


Image acquisition
Image pre-processing
Main image segmentation
Segmented image post-processing
Image segmentation
Feature extraction
Feature-based classification
Image understanding
Application-specific feature extraction


Input image
Pre processing
Laplacian and sobel gradient
Clustering
Morphological process
Segment text
Data set
Classification
Result

Chapter 1: Digital Image Formation And Low Level Processing

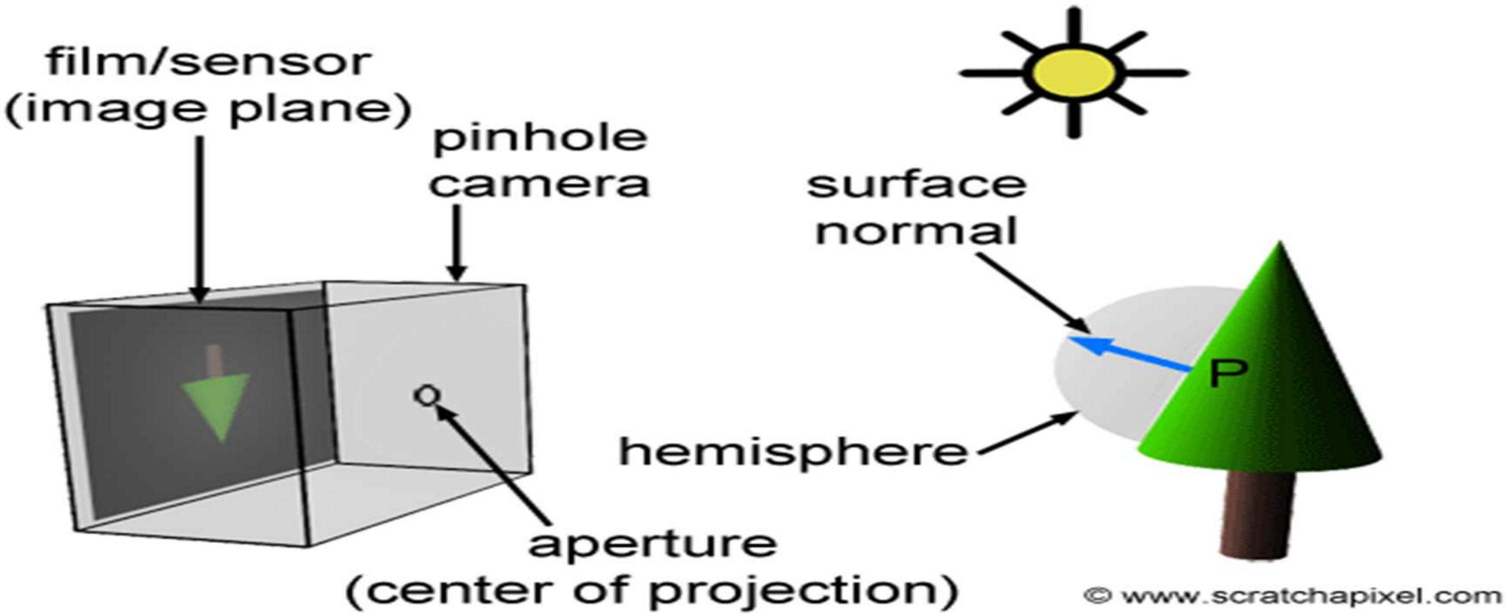
Fundamentals of Image Formation

Image: Projection of 3D Scene to 2D plane. We need to understand the geometric and photometric relation between the scene and its image.

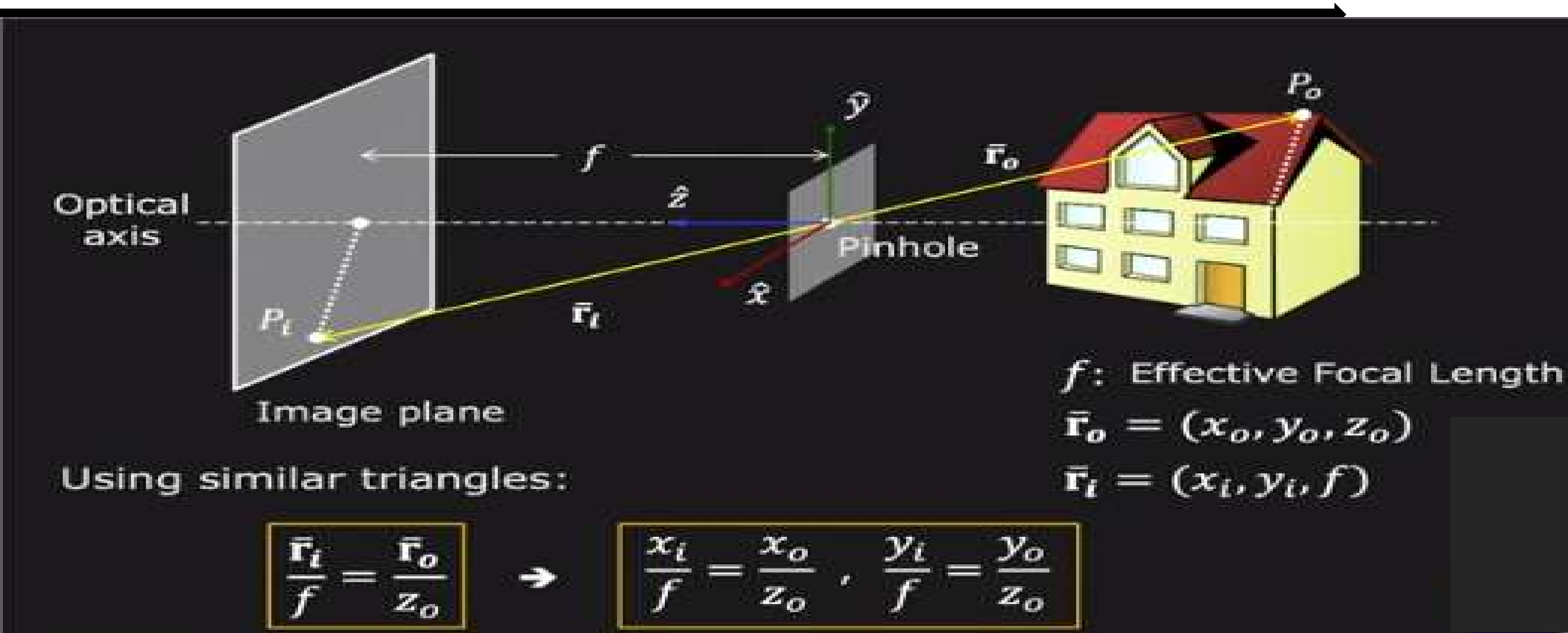
Topics:

- Pinhole and Perspective projection
- Image formation using lenses
- Lens related issues
- Wide angle cameras
- Biological eyes

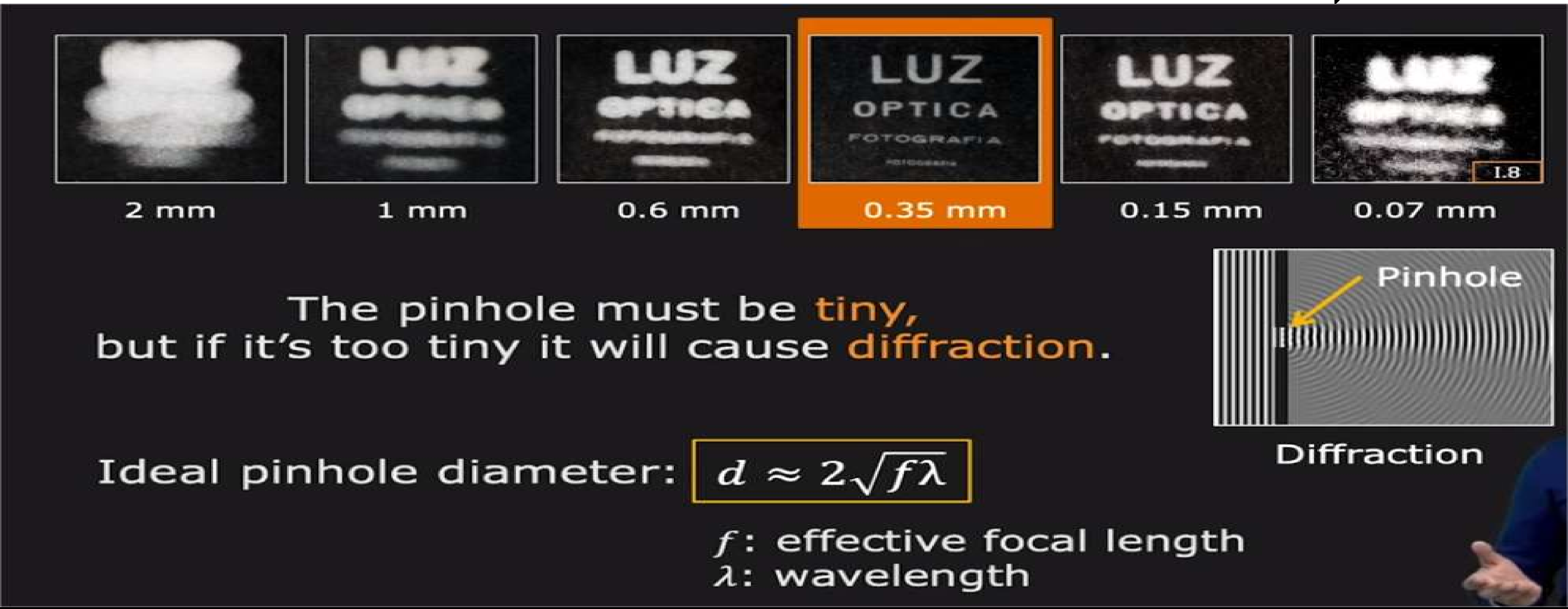
Perspective Projection with pinhole



Perspective Projection



Ideal size of a pinhole camera



Vanishing Point

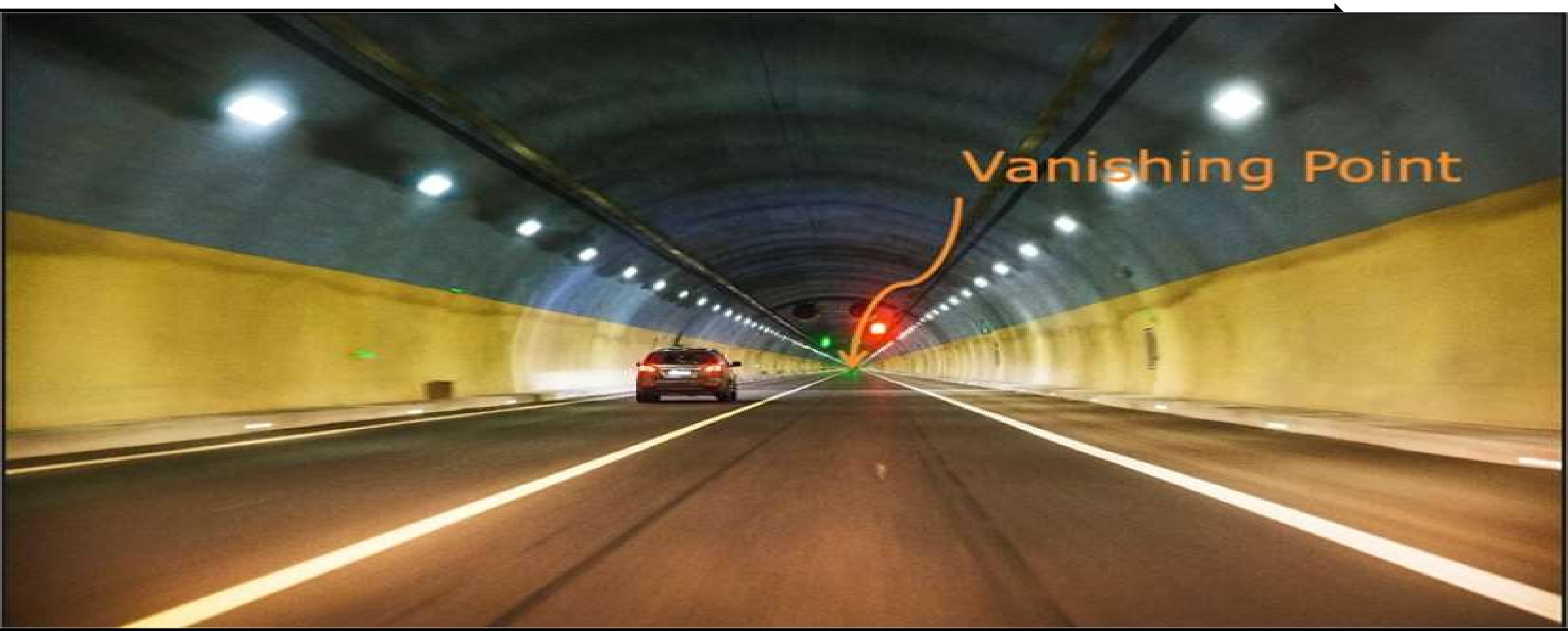
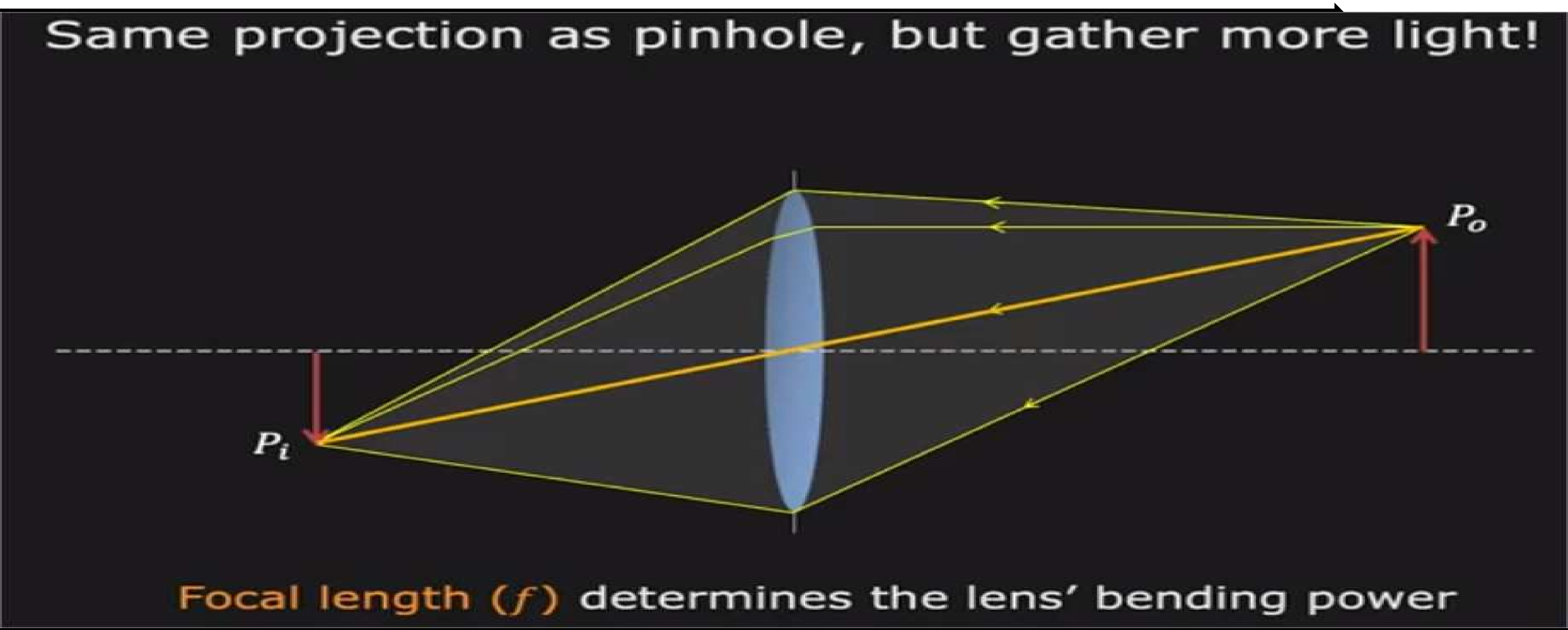
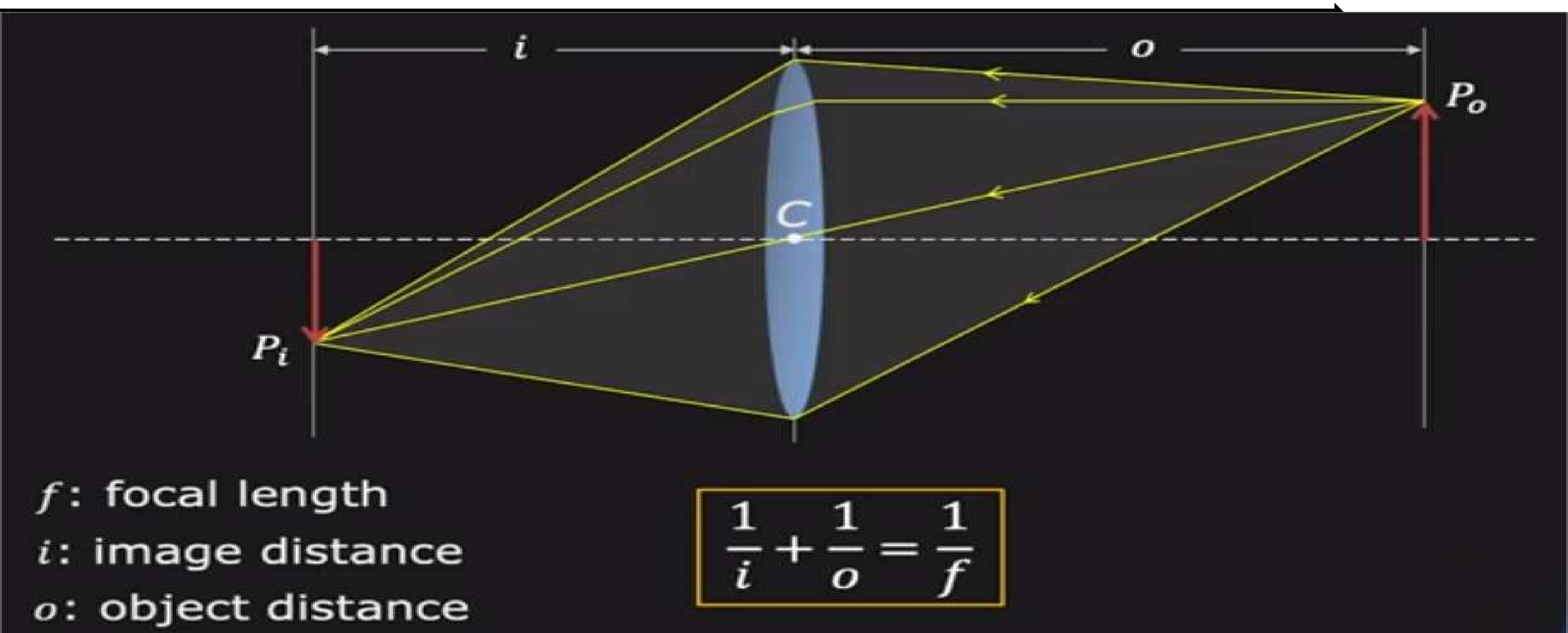


Image Formation using Lenses

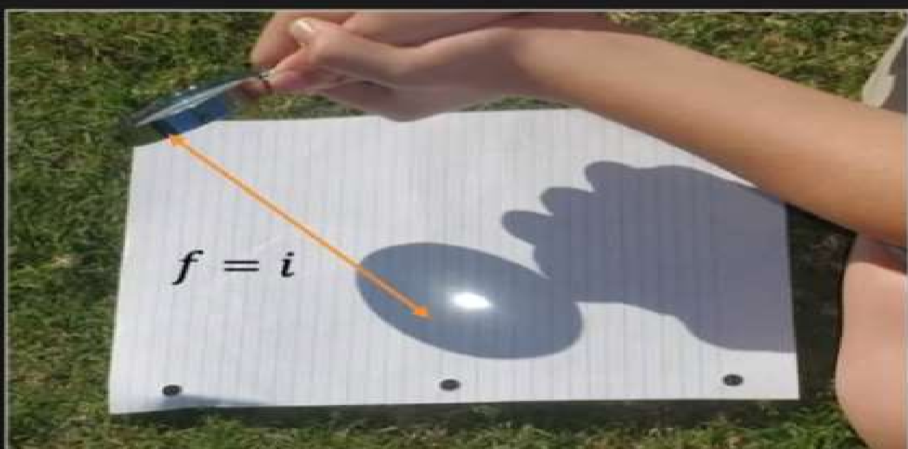


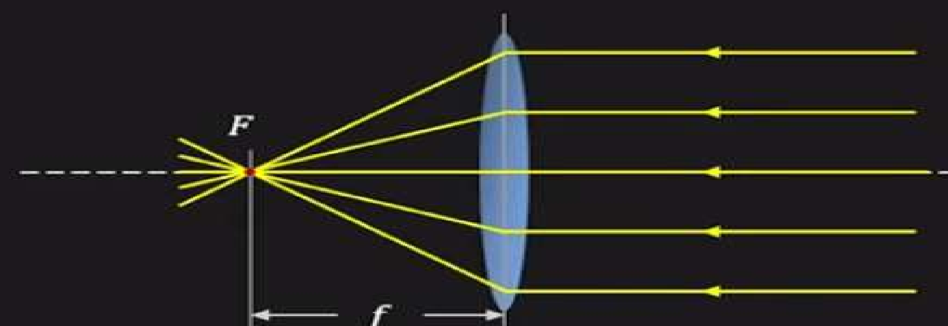
Gaussian lens law



How do you find the focal length

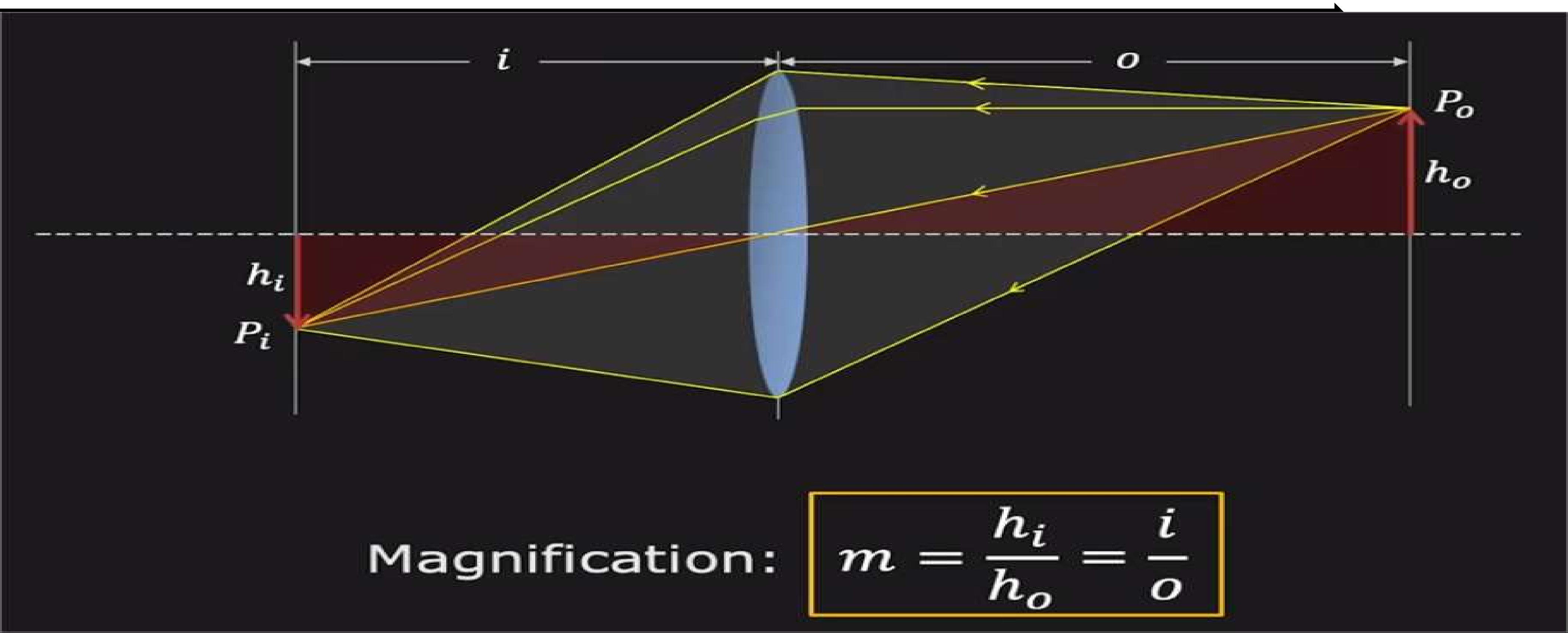
$$\frac{1}{i}+\frac{1}{o}=\frac{1}{f} \quad \Rightarrow \quad \text{If } o=\infty\text{, then } f=i$$



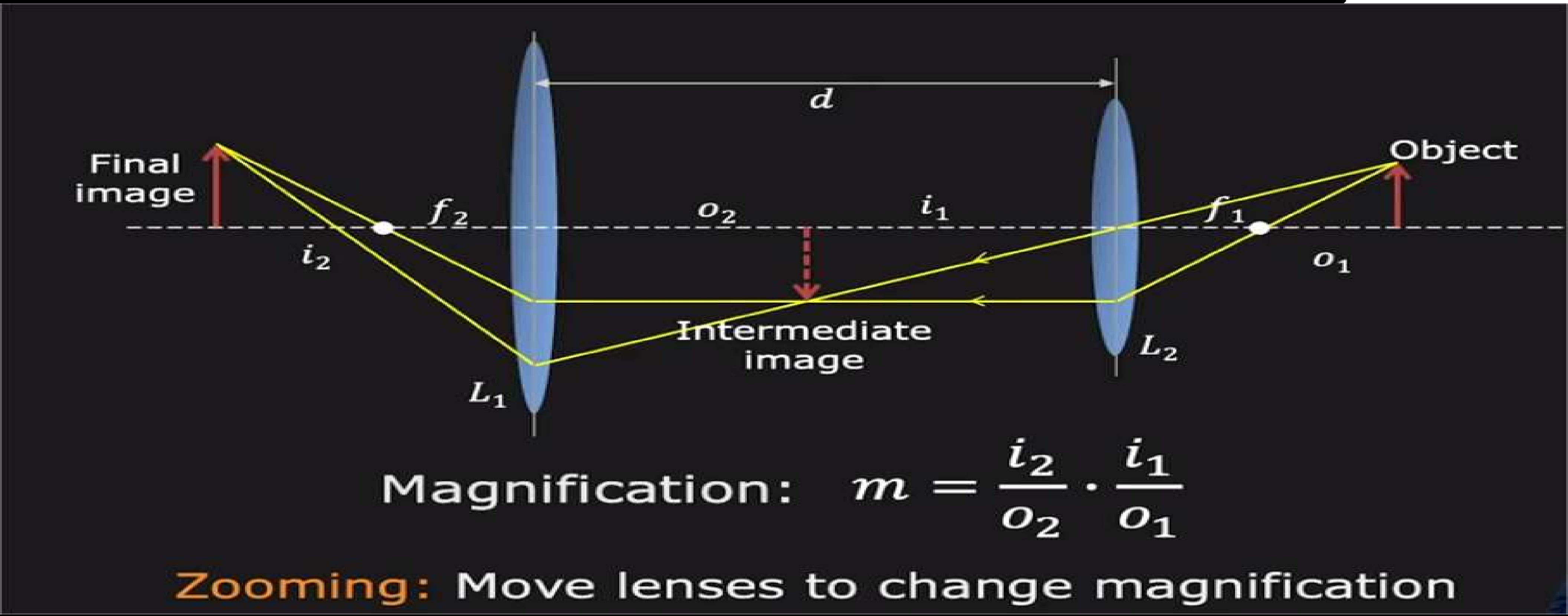


Focal length: Distance at which incoming rays that are parallel to the optical axis converge.

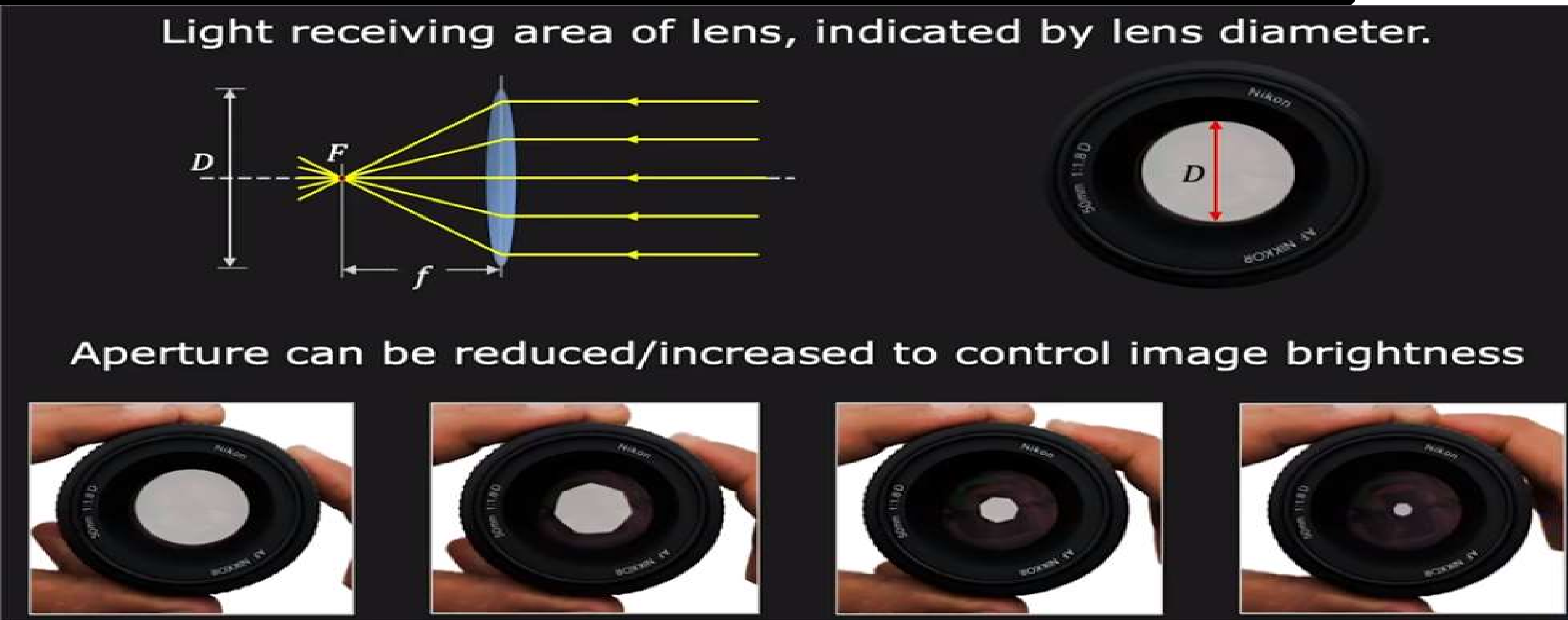
Image Magnification



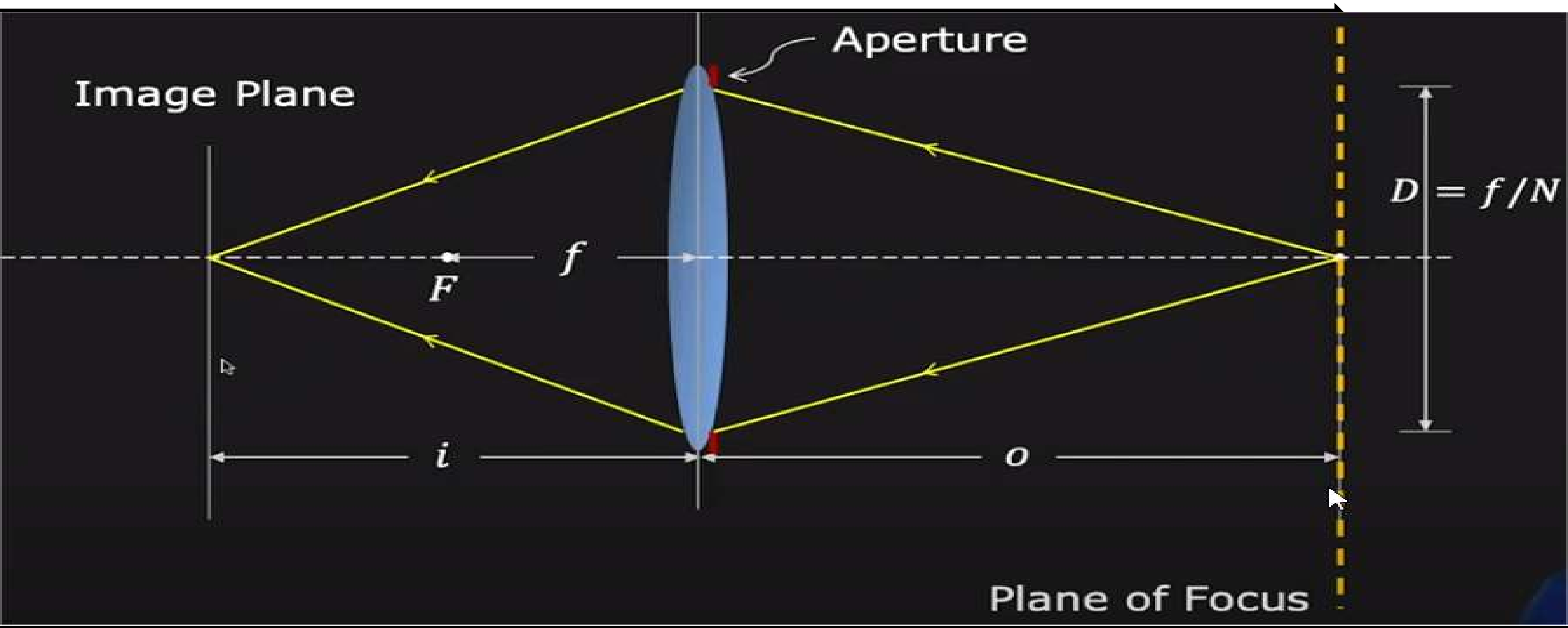
Two lenses System

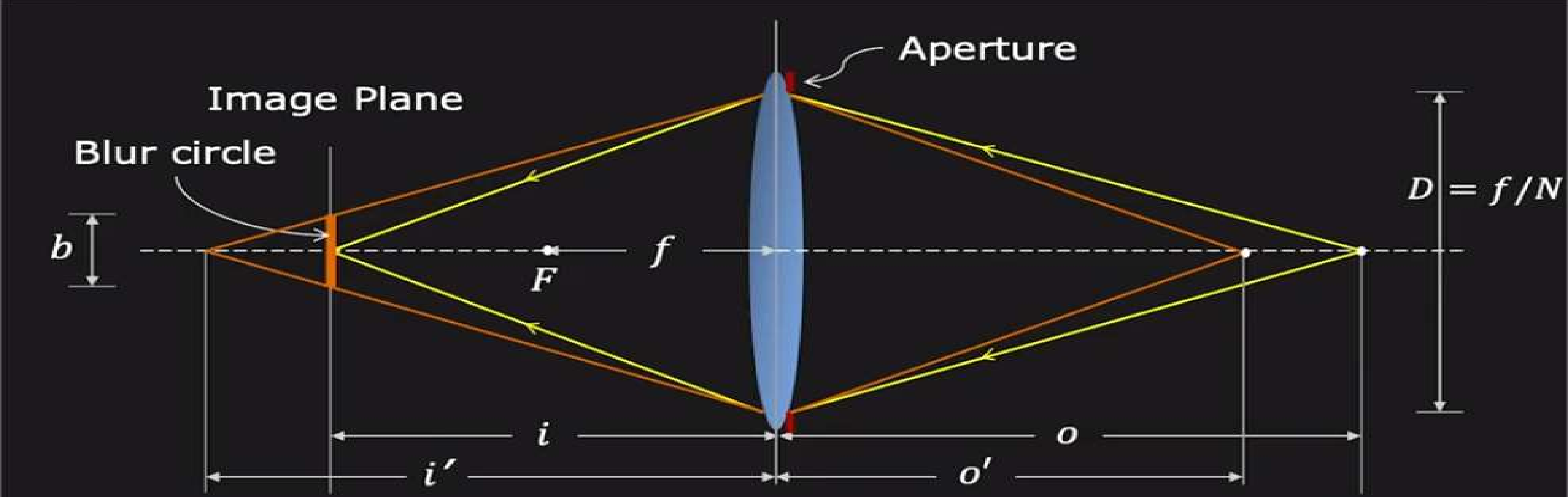


Aperture


Light receiving area of lens, indicated by lens diameter.
D
F
f
D
Aperture can be reduced/increased to control image brightness

Lens Defocus





From similar triangles:

$$\frac{b}{D} = \frac{|i' - i|}{i'}$$

Blur circle diameter:

$$\boxed{b = \frac{D}{i'}|i' - i|} \quad \boxed{b \propto D \propto \frac{1}{N}}$$

Focused Point

$$\frac{1}{i} + \frac{1}{o} = \frac{1}{f}$$

$$i = \frac{of}{o - f}$$

Defocused Point

$$\frac{1}{i'} + \frac{1}{o'} = \frac{1}{f}$$ (Gaussian Lens Law)

$$i' = \frac{o'f}{o' - f}$$

$$i' - i = \frac{f}{(o' - f)} \cdot \frac{f}{(o - f)} \cdot (o - o')$$

$$b = Df \left| \frac{(o - o')}{o'(o - f)} \right|$$

$$\boxed{b = \frac{f^2}{N} \left| \frac{(o - o')}{o'(o - f)} \right|}$$

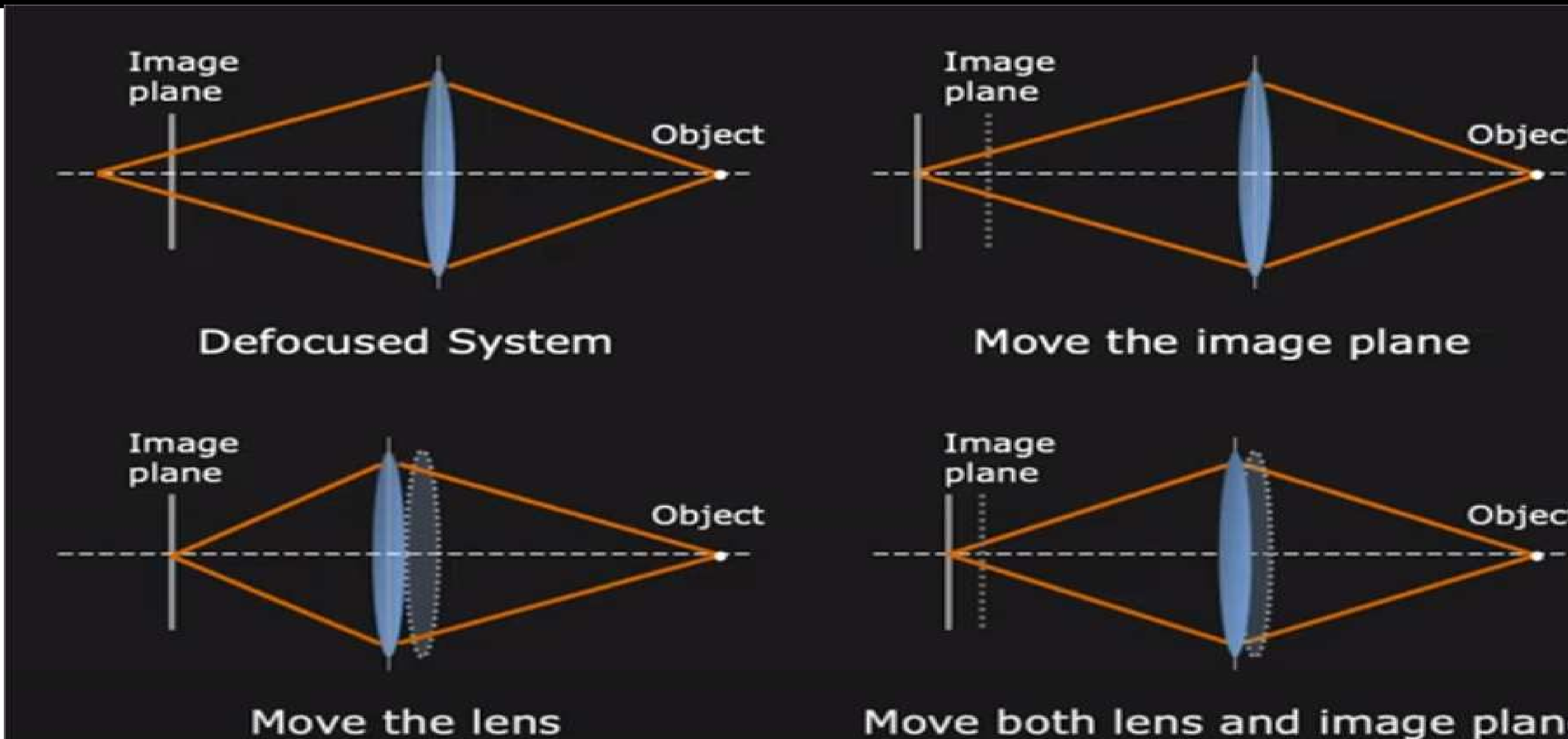
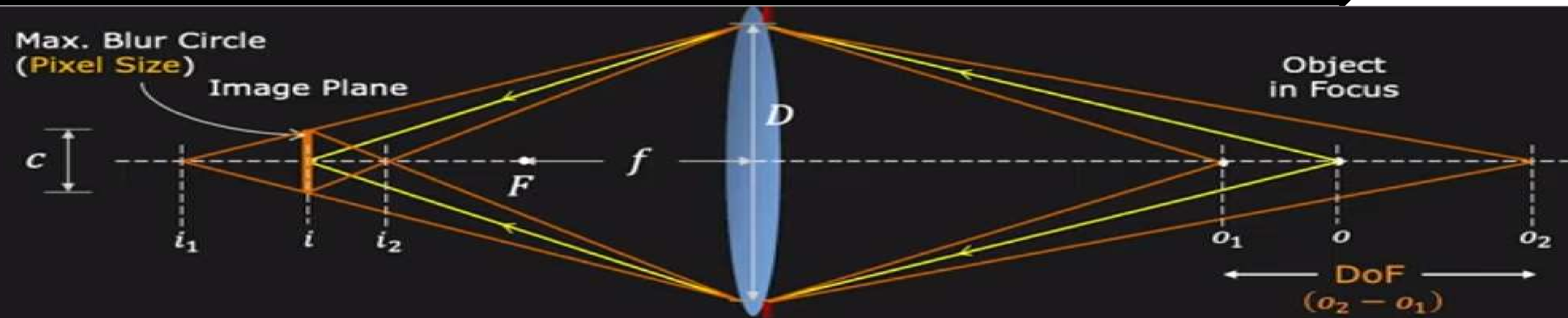


Image
plane
Object
Defocused System
Image
plane
Objec
Move the image plane
Image
plane
Object
Move the lens
Image
plane
Objec
Move both lens and image plan

Depth of Field

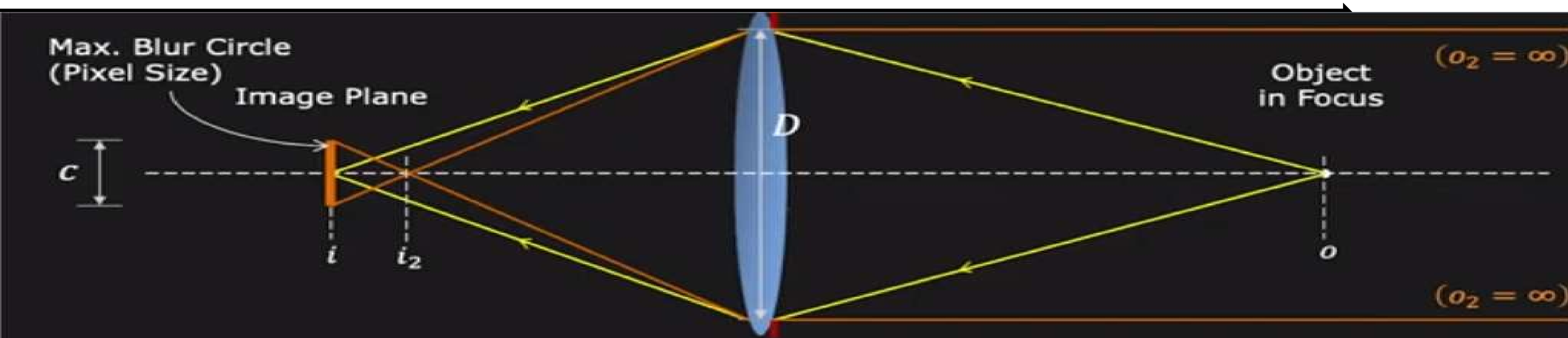


If o_1 and o_2 are the nearest and farthest distances respectively for which blur circle is maximum c, then:

$$c = \frac{f^2(o - o_1)}{N o_1 (o - f)} \qquad c = \frac{f^2(o_2 - o)}{N o_2 (o - f)}$$

Depth of Field: $$\boxed{o_2 - o_1 = \frac{2 o f^2 c N (o - f)}{f^4 - c^2 N^2 (o - f)^2}}$$

Hyper Focal distance



The closest distance $o = h$ the lens must be focused to keep objects at infinity $(o_2 = \infty)$ acceptably sharp (blur circle $\leq c$).

Hyperfocal Distance: $\boxed{h = \frac{f^2}{Nc} + f}$