COMPUTER VISION

Subject Code: 3171614

Prepared By:

Prof. Janki Patel

Department of IT

SPCE, Bakrol

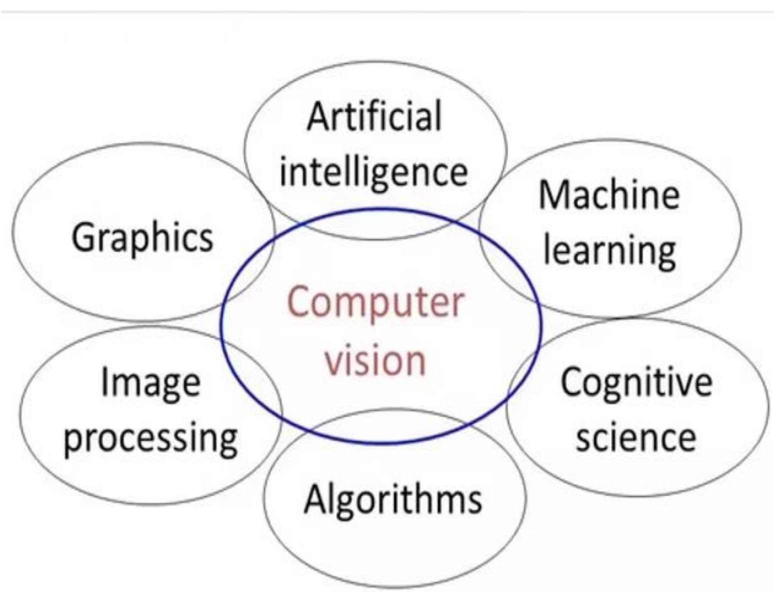
Unit:1
Overview of computer vision and its applications

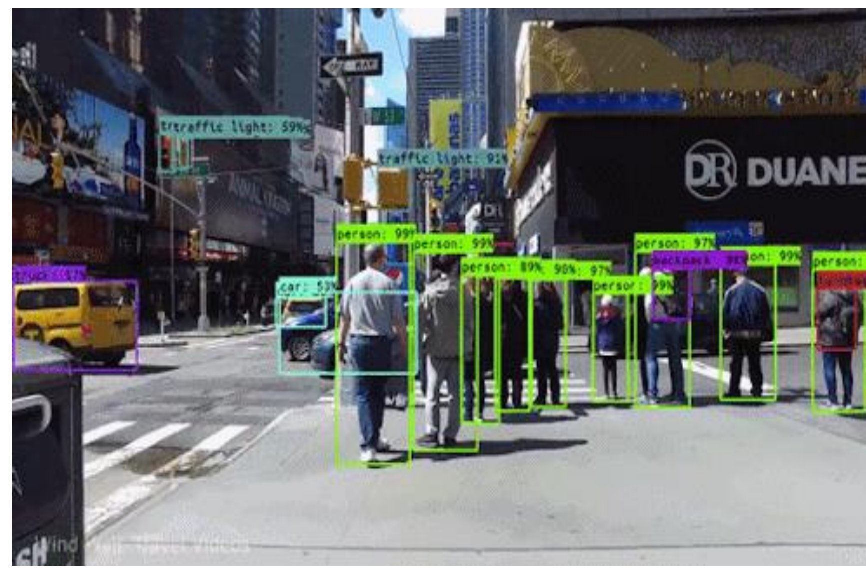
Content

- **Image Formation and Representation:**
- Imaging geometry,
- radiometry,
- digitization,
- cameras and Projections,
- rigid and affine transformation

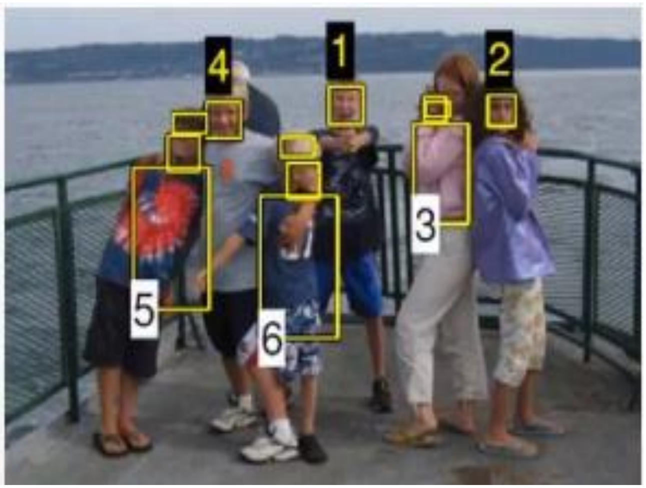
What is Computer Vision?

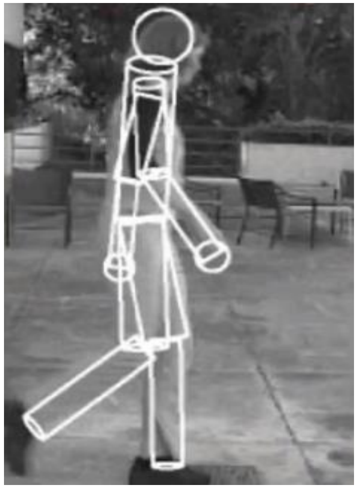
- Computer vision is the field of artificial intelligence and computer science that aims at giving computer a visual understanding of the world , and is the heart of Hayo’s powerful algorithms.
- It is one of the main components of machine understanding:

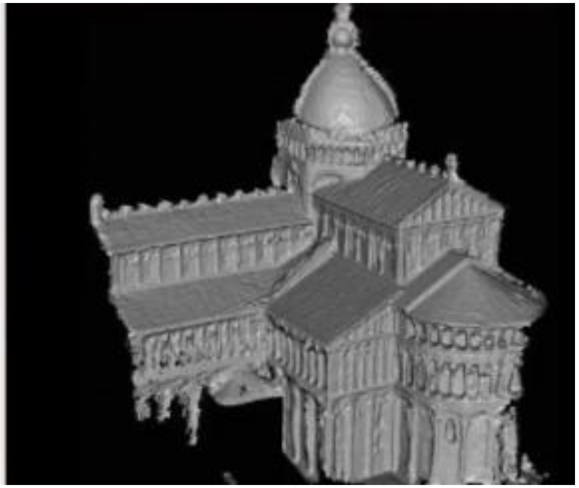




Example of computer vision



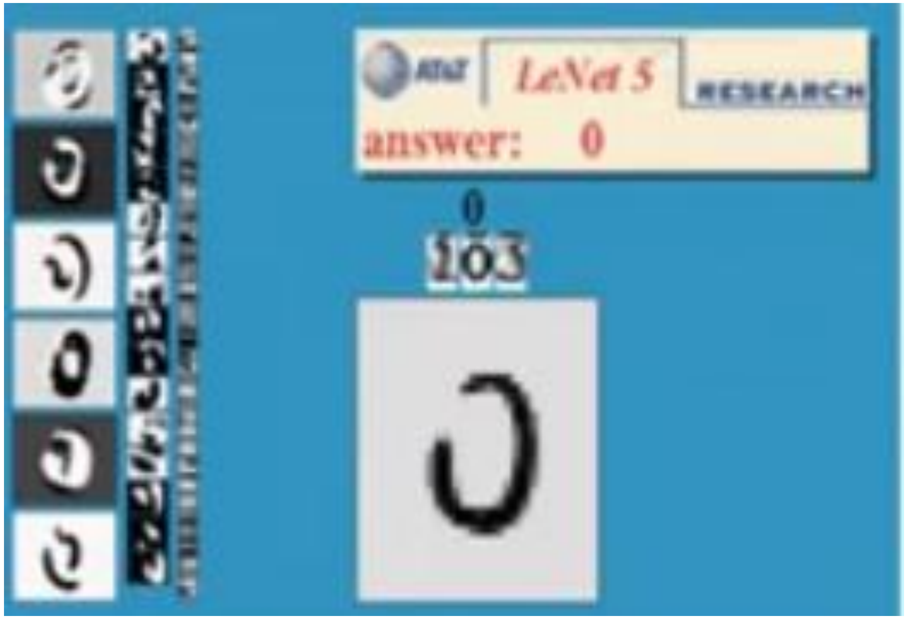




Application of computer vision

- The good news is that computer vision is being used today in a wide variety of real-world application, which include:
- Optical Character recognition(OCR)
- Machine Inspection
- Retail
- 3D model building
- Medical imaging
- Automation safety
- Match move

- **Optical Character recognition(OCR):** Reading handwritten code postal codes on letters and automatic number plate recognition[ANPR].

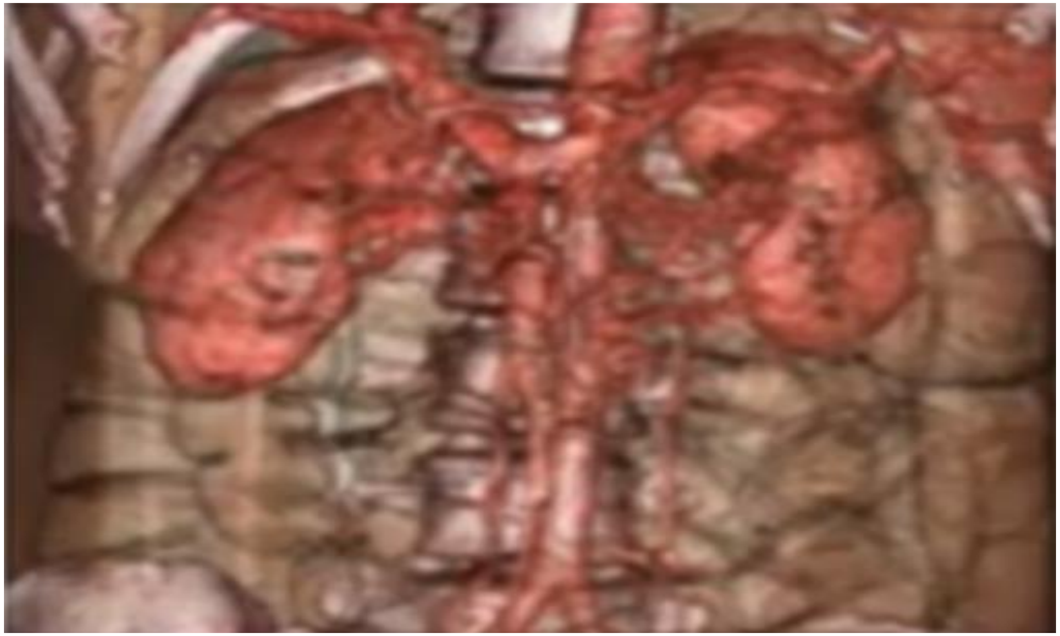


- **Machine Inspection:** Rapid parts inspection for quality assurance using stereo vision with specialized illumination to measure tolerances on aircraft wings or auto body parts of looking for defects in steel casting using X-ray vision.

- **Retail:** Object recognition for automated checkout lanes.

- **3D model building:** fully automated construction of 3D models from aerial photographs used in system such Bing Maps.

- **Medical imaging:** Registering pre-operative and intra-operative imagery or performing long-term studies of people’s brain morphology as they age.



- **Automotive safety:** Detecting unexpected obstacles such as pedestrians on the street, under conditions where active vision techniques such as radar or lidar do not work well.

- **Match move:** Merging computer imagery(CGI)with live action footage by tracking feature points in the source video to estimate the 3D camera motion and shape of environment. Such techniques are widely used in Hollywood(e.g, in movies such as Jurassic Park)they also require the use of precise matting to insert new element between foreground and background elements

- **Motion Capture:** Using retro-reflective markers viewed from multiple cameras or other vision-based techniques to capture actors for computer animation
- **Surveillance:** monitoring for intruders, analyzing highway traffic and monitoring pools for drowning victims

- Now, we focus more on broader consumer level applications, such as fun things you can do with your own personal photographs and video. These include:

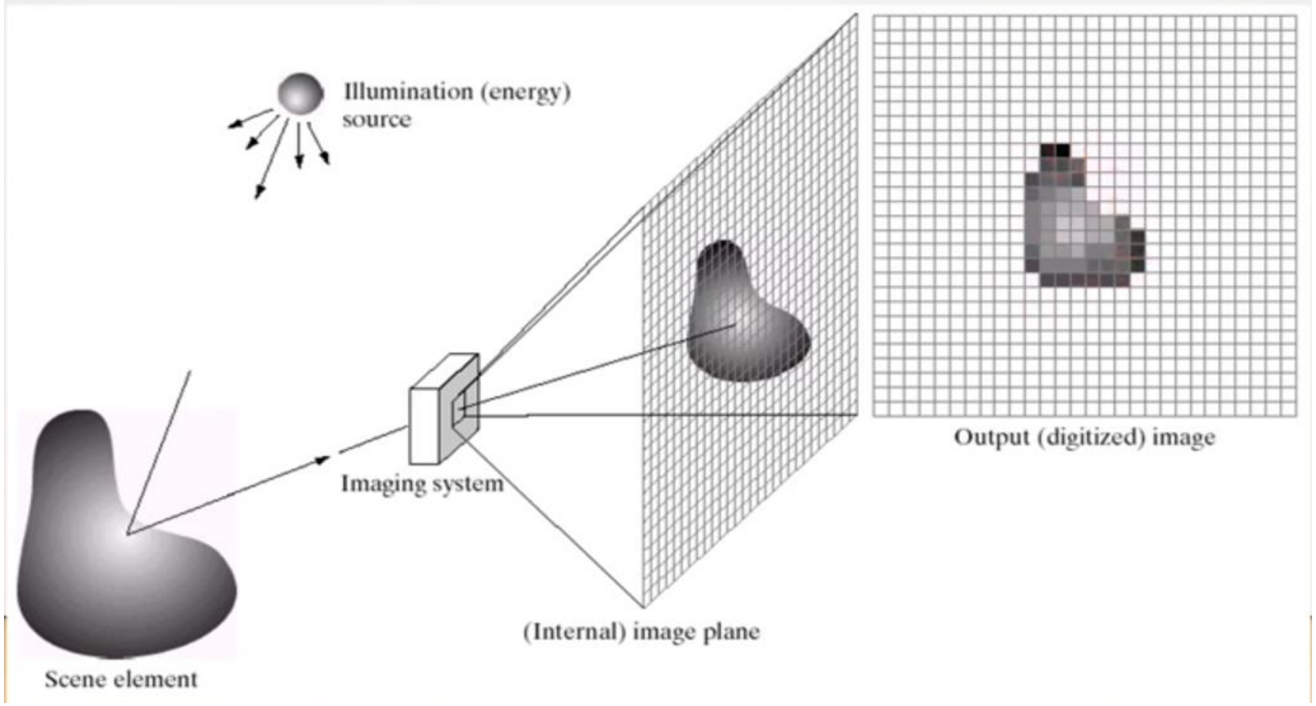
- **Stitching:** turning overlapping photos into a single seamlessly stitched panorama

- **Exposure bracketing:** Merging multiple exposures taken under challenging lighting conditions into a single perfectly exposed image.

Image Digitization

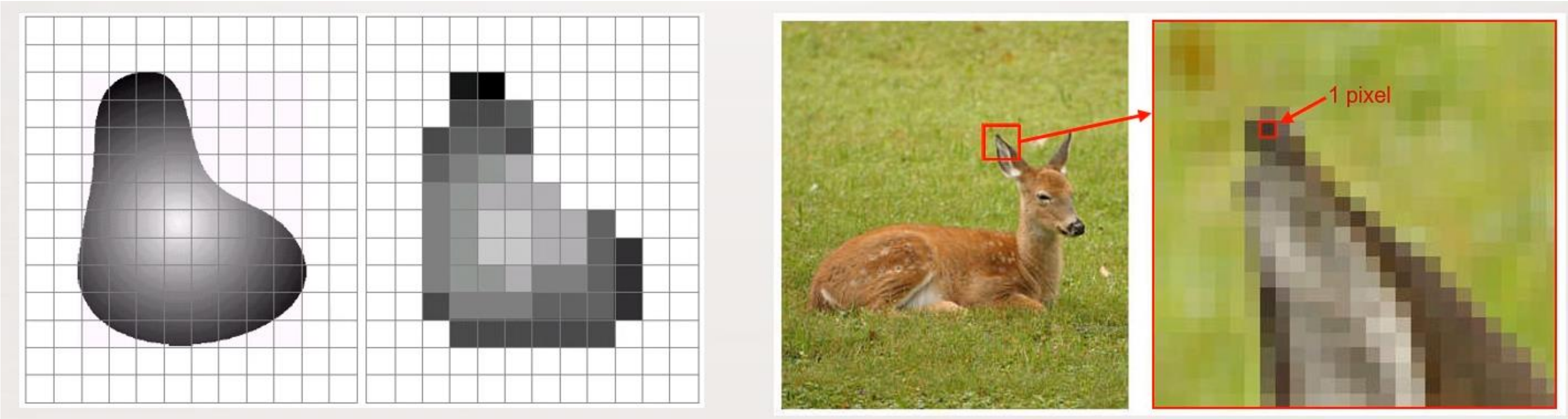
What is a Digital Image?

- A digital image is a representation of a two-dimensional image as a finite set of digital values, called picture elements or pixels.



Cont…

- Pixel values typically represent gray levels, colores, heights, opacities etc.
- Remember digitization implies that a digital image is an approximation of a real scene.



Cont…

- Common image formats include:
- 1 sample per point(B&W or Grayscale)
- 3sample per point(Red, Green, and Blue)
- sample per point(Red, Green, and Blue, and “Alpha”,a.k.a. Opacity)



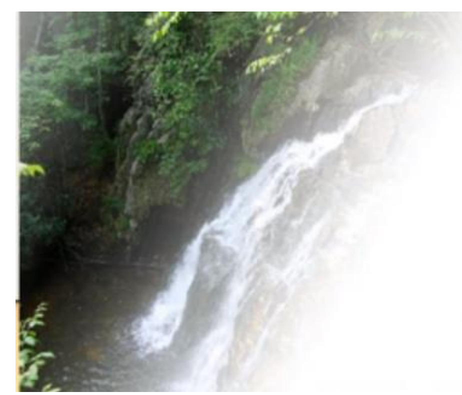


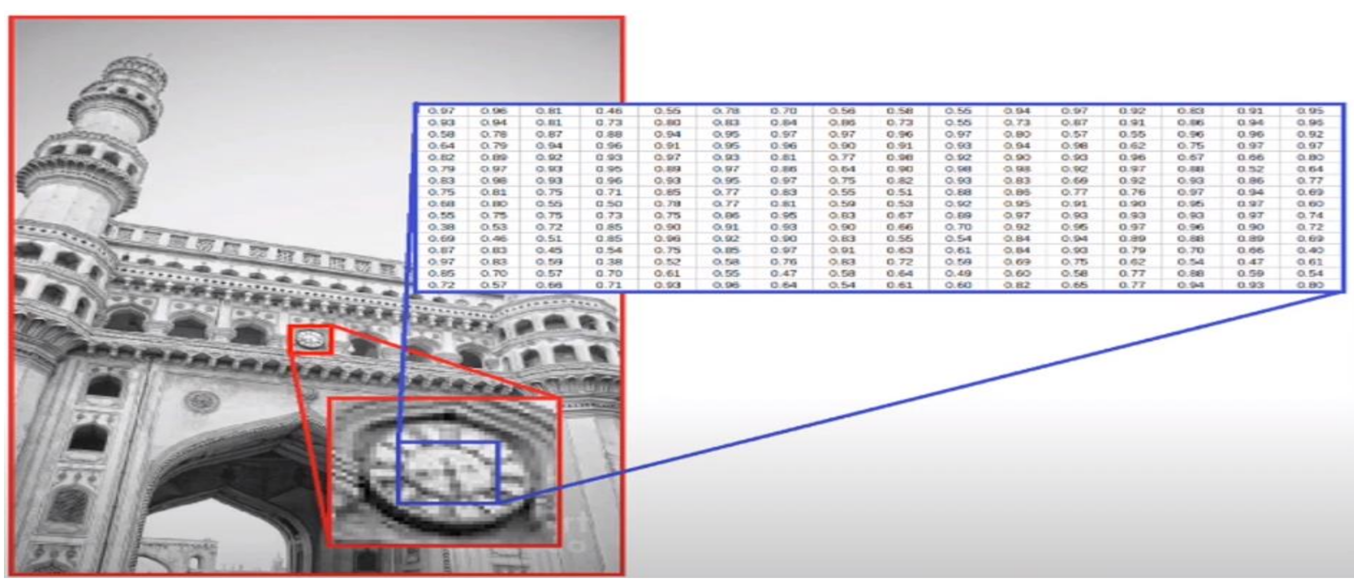
Image Representation

Image Representation

- After getting an image ,it is important to devise ways to represented. Let’s look at the most common ways to represent an image.

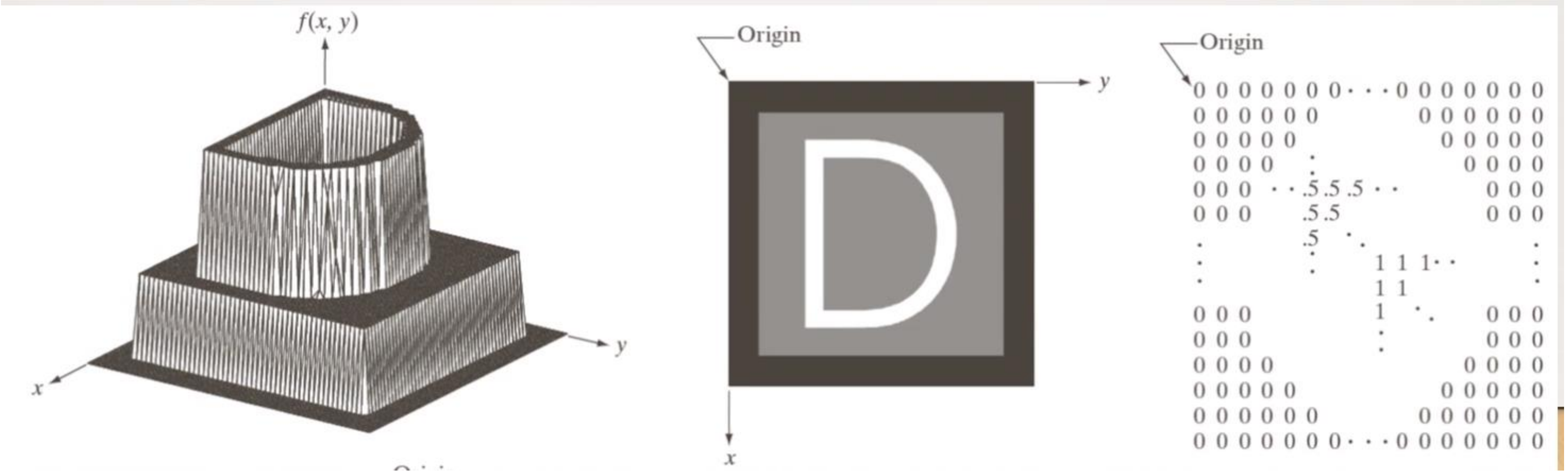
Image Representation

- **Image as matrix**
- The simplest way to represent the image is in the form of a matrix.
- In figure we can see that a part of the image, i.e., the clock, has been represented as a matrix. A similar matrix will represent the rest of the image too.



Cont…

- It is commonly seen that people use up to a byte to represent every pixel of the image. This means that values between 0 to 255 represent the intensity for each pixel in the image where 0 is black and 255 is white. For every color channel in the image, one such matrix is generated.



Cont…

- **Image as a function**
- An image can also be represented as a function. An image(grayscale)can be thought of as a function that takes in a pixel coordinate and gives the intensity at that pixel.
- It can be written as function $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ that outputs the intensity at any input point(x, y).The value of intensity can be between 0 to 255 or 0 to 1 if values are normalized.

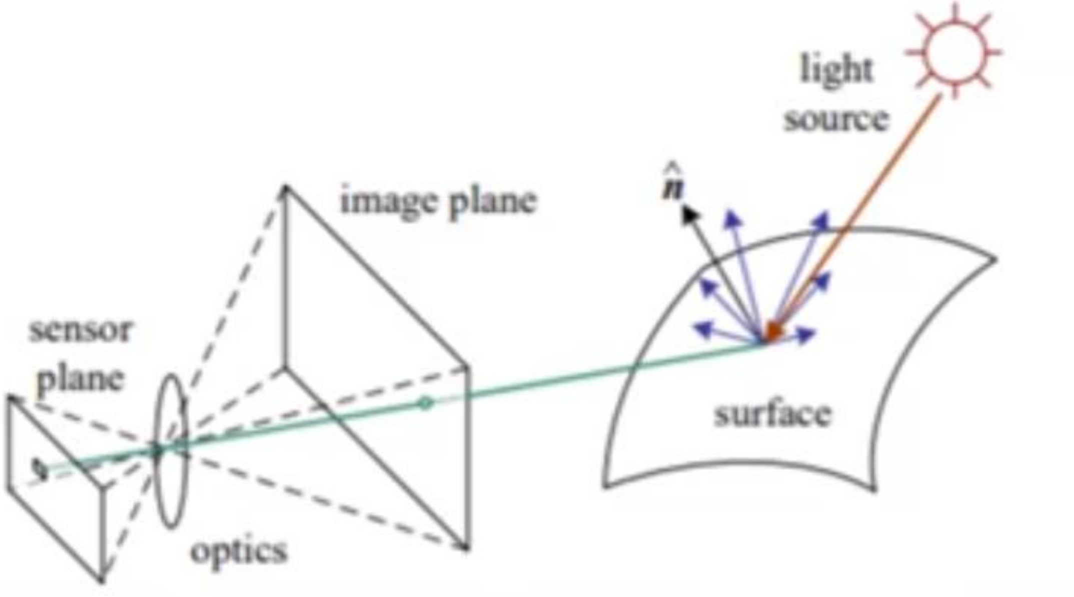
Image Formation

Image Formation

- In modeling any image formation process, geometric primitives and transformations are crucial to project 3-D geometric features. However, apart from geometric features, image formation also depends on discrete color and intensity values.
- It needs to know the lighting of the environment, camera optics, sensor properties, etc. Therefore, while about image formation in computer vision.
    - (200,100,50):RGB
    - (122):Gray scale

Photometric Image Formation

- In figure a simple explanation of image formation. The light from a source is reflected on a particular surface. A part of that reflected light goes through an image plane that reaches a sensor plane via optics.

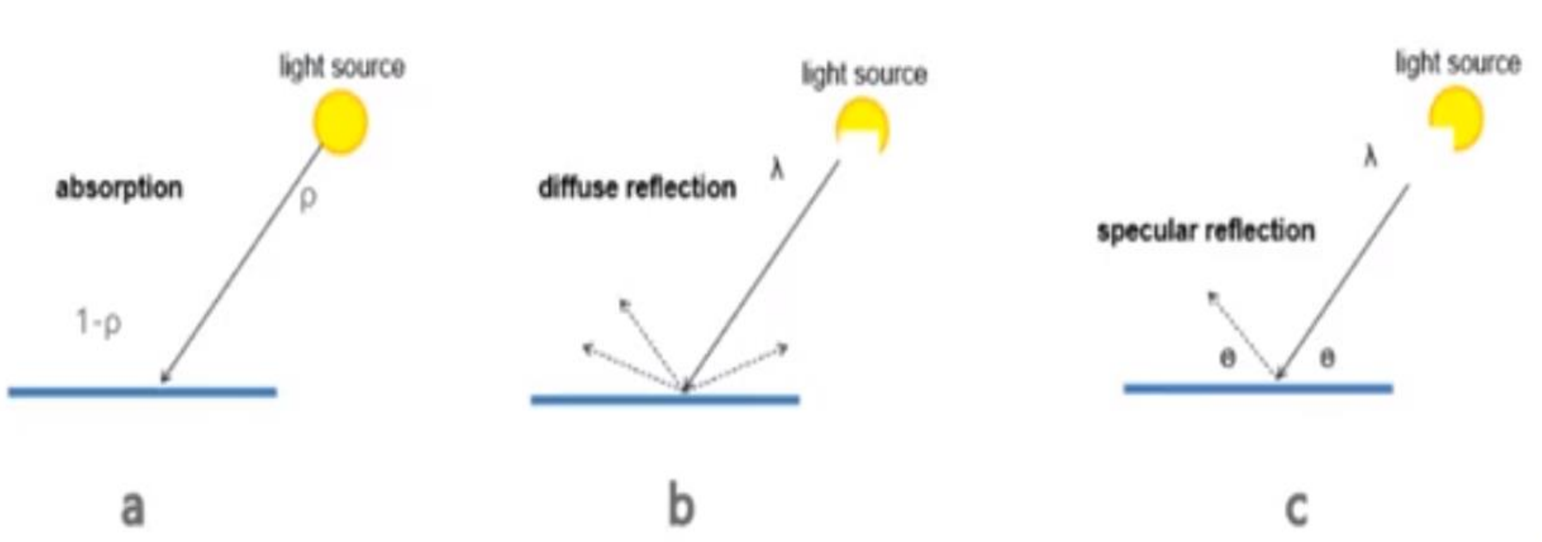


Cont…

- **Some factors that affect image formation are:**
- The strength and direction of the lightemitted from the source.
- The material and surface geometry along with other nearby surfaces.
- Sensor capture properties

Cont…

- **Reflection and Scattering**
- Image cannot exist without light. Light sources can be a point or an area light source. When light hits a surface, three major reactions might occur.



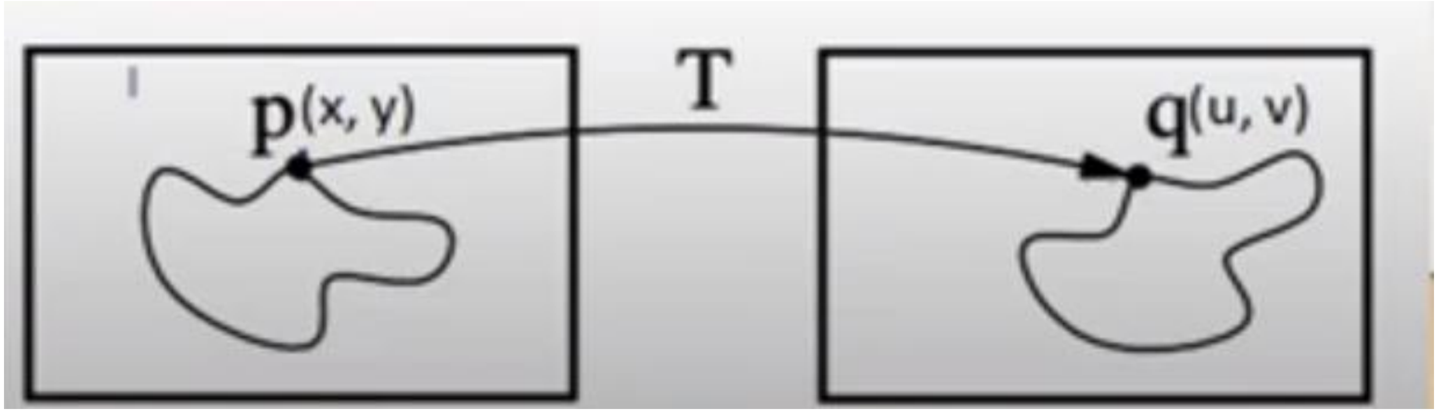
Cont…

- **Color**
- From a viewpoint of color, we know visible light is only a small portion of a large electromagnetic spectrum.
- Two factors are noticed when a colored light arrives at a sensor:
  1. Color of the light
  2. Color of the surface

Image Geometric/Spatial Transformation

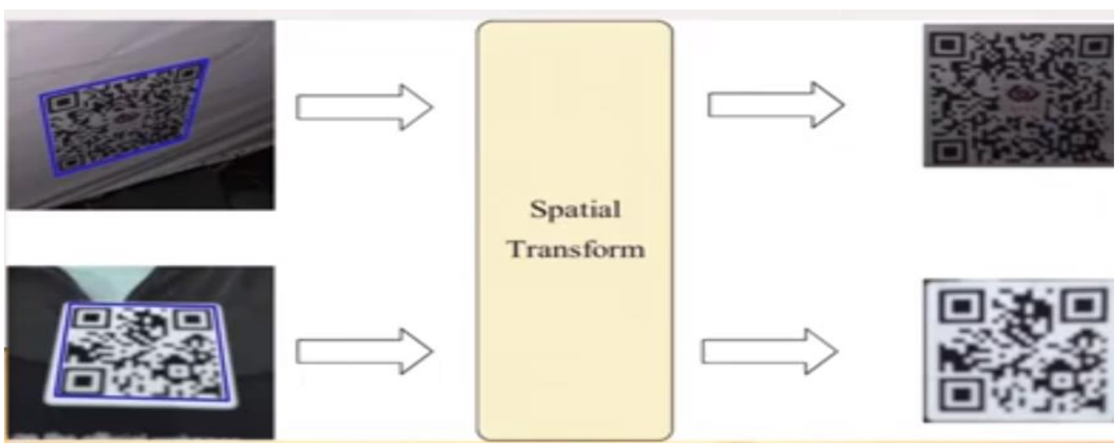
Image Geometric/Spatial Transformation

- Image geometric that means changing the geometry of an image.
- Geometric transforms permit the elimination of distortion that occurs when an image is captured.
- A spatial transformation of an image is a geometric transformation of the image coordinate system.
- In spatial transformation each point(x,y) of image A is mapped to a point(u,v) in a new coordinate system.



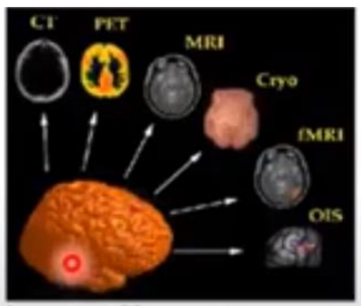
Why it is used?

- Some person clicking the pictures of the same place at different times of the day and year to visualize the changes. Every time he clicks the picture, it’s not necessary that he clicks the picture at the exact same angle. So for better visualization, he can align all the images at the same angle using geometric transformation.

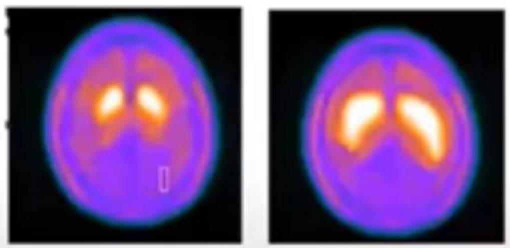


Why geometric transformation in required?

- Image registration is the process of transforming different sets of data into one coordinate system.



from different sensors


at different time

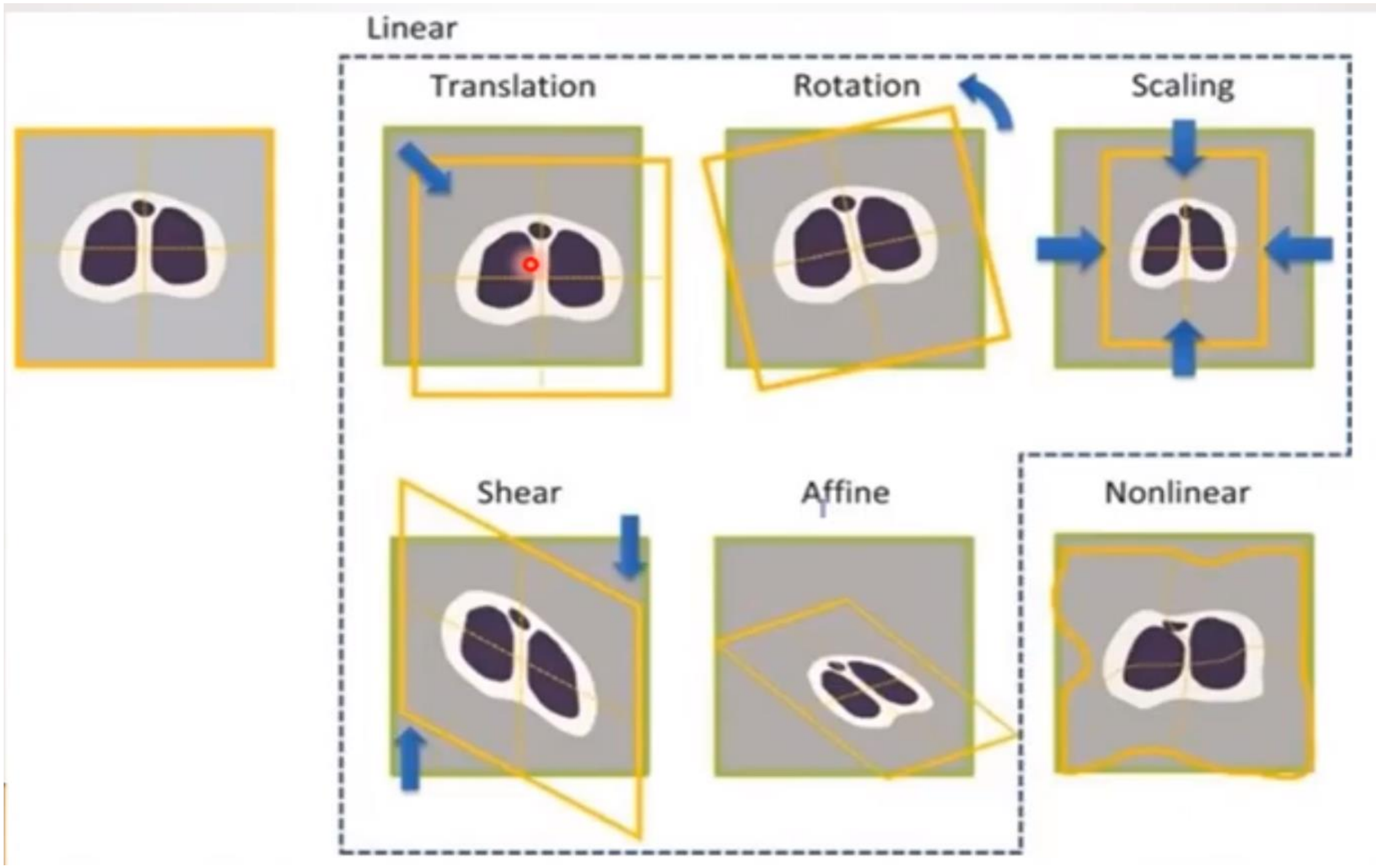

from different depths


from different viewpoints


from different viewpoints

Types of geometric Transformation

Types of geometric Transformation



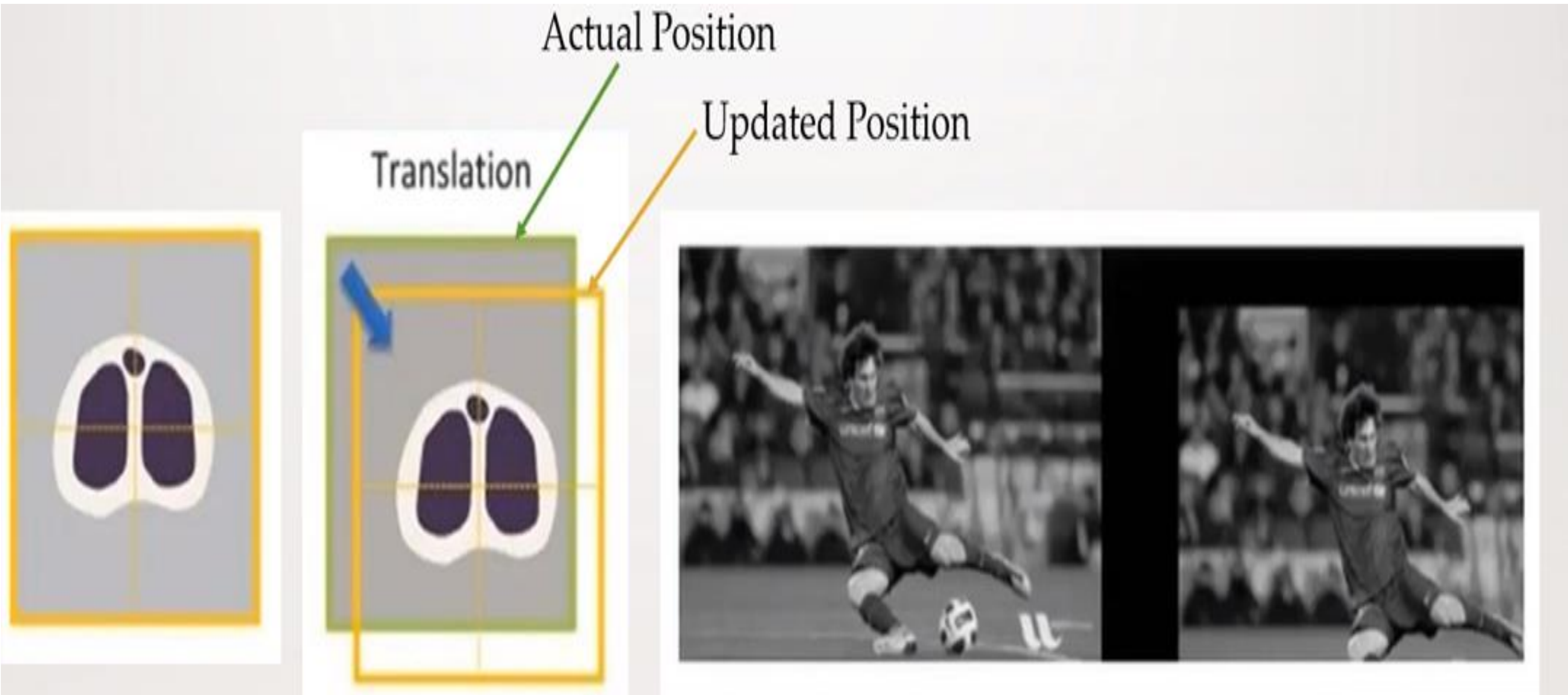
1. Translation

- Translation is the shifting of the object's location. If you know the shift in(x,y) direction, let it be, you can create the transformation matrix as follows:

$$M = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix}$$

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Cont…



2.Rotation

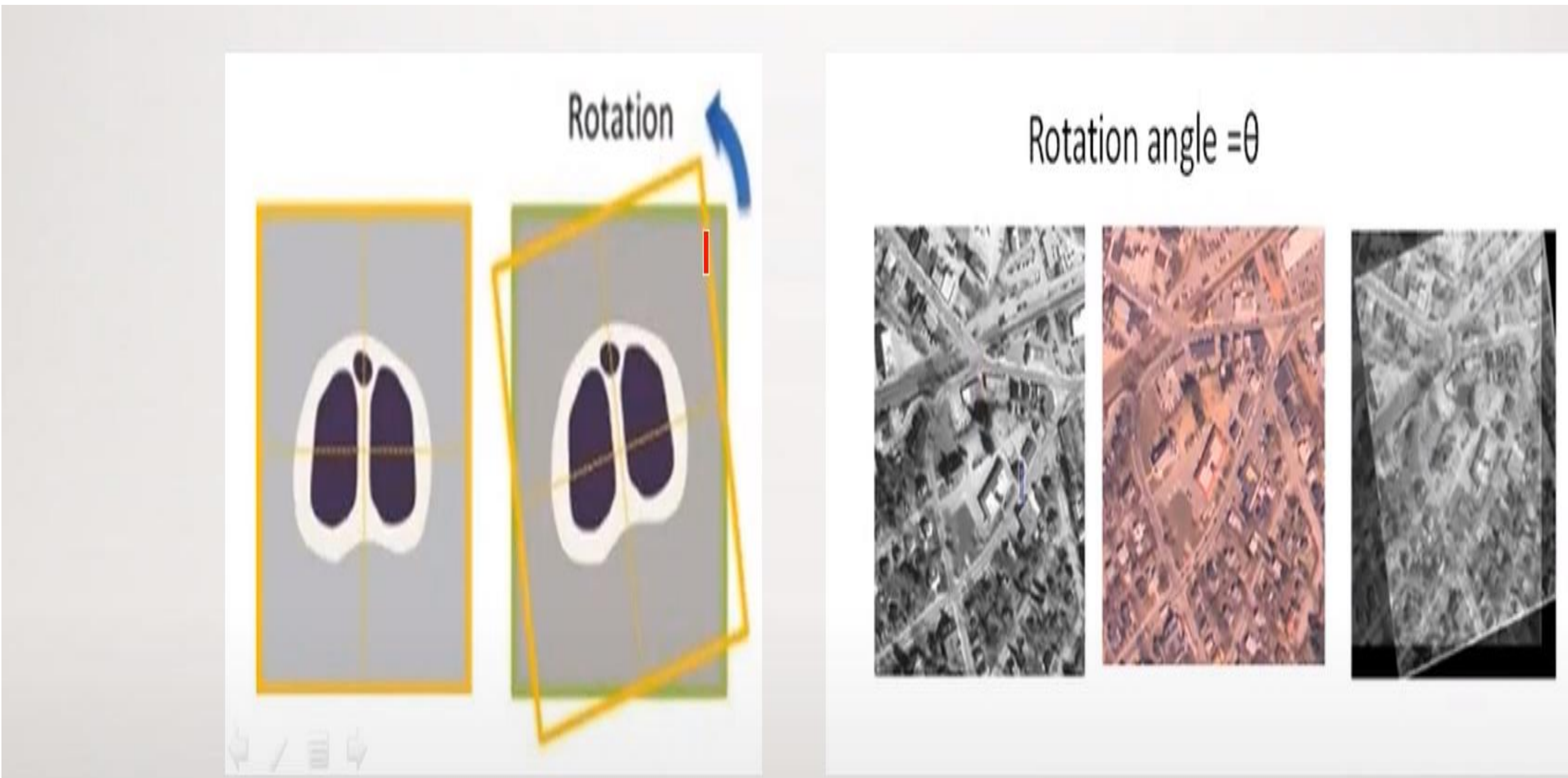
- This technique rotates an image by a specified angle and by the given axis or point.
- The points that lie outside the boundary of an output image are ignored.
- Rotation about the origin by an angle θ is given by:

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$$u = x\cos\theta + y\sin\theta$$

$$v = -x\sin\theta + y\cos\theta$$

Cont…



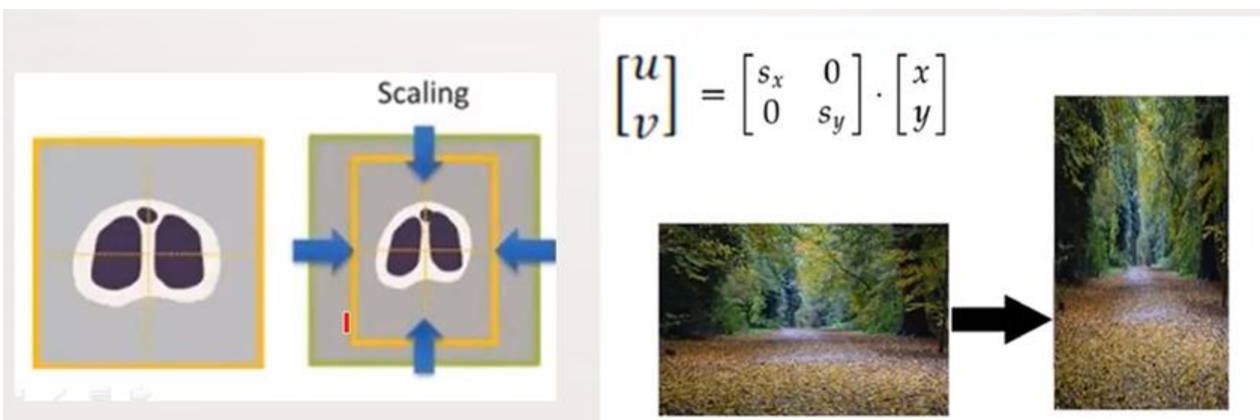
3.Scaling

- Scaling means resizing an image which means an image is made bigger or smaller in x/y direction.
- We can resize an image in terms of scaling factor.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} s_x & 0 \\ 0 & s_y \end{bmatrix} \cdot \begin{bmatrix} x \\ y \end{bmatrix}$$

Cont…

If we have an image of size (300 x 400) and we want to transform it into an image of shap
(600 x 200).
The scaling in x- direction will be : 600/300 = 2. (we denote it as Sx = 2)
Similarly Sy = 200/400 = 1/2.

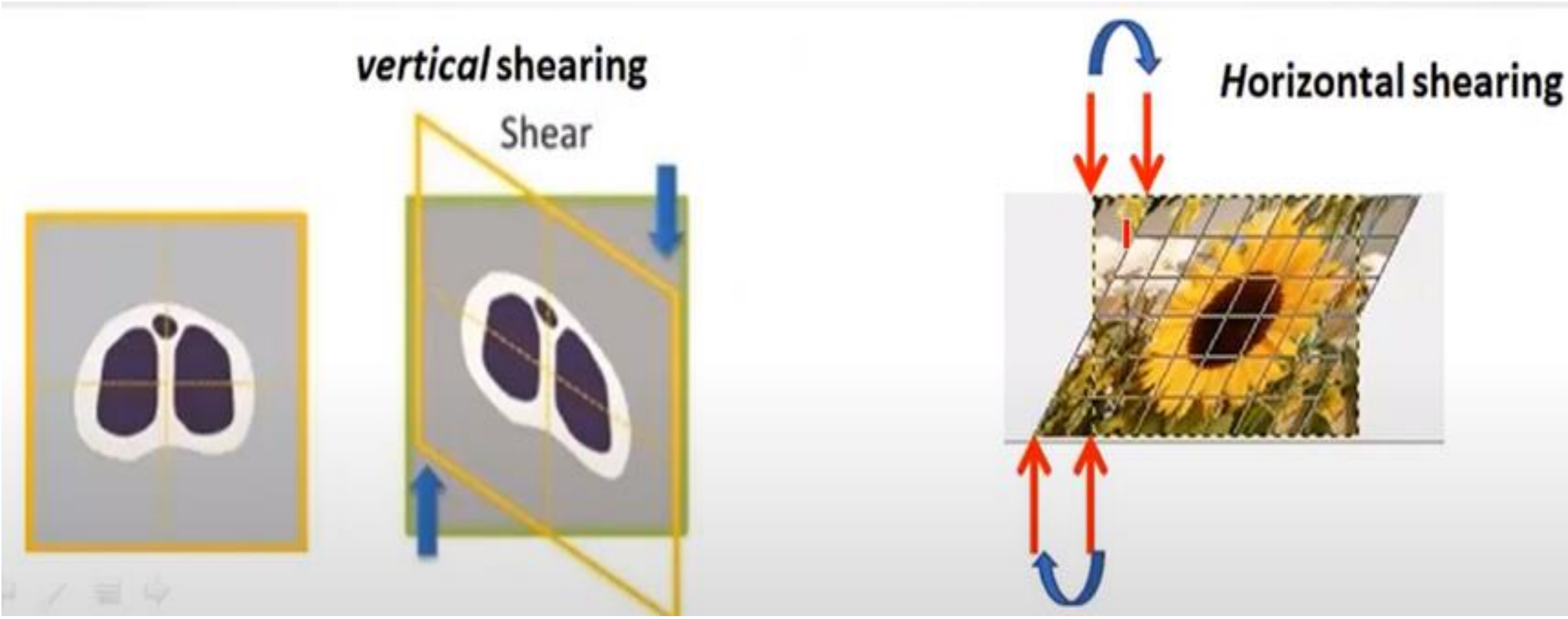


- (x,y)=(300,400)
- 300 row and 400 col.
- 300=height 400=width

4.Shearing

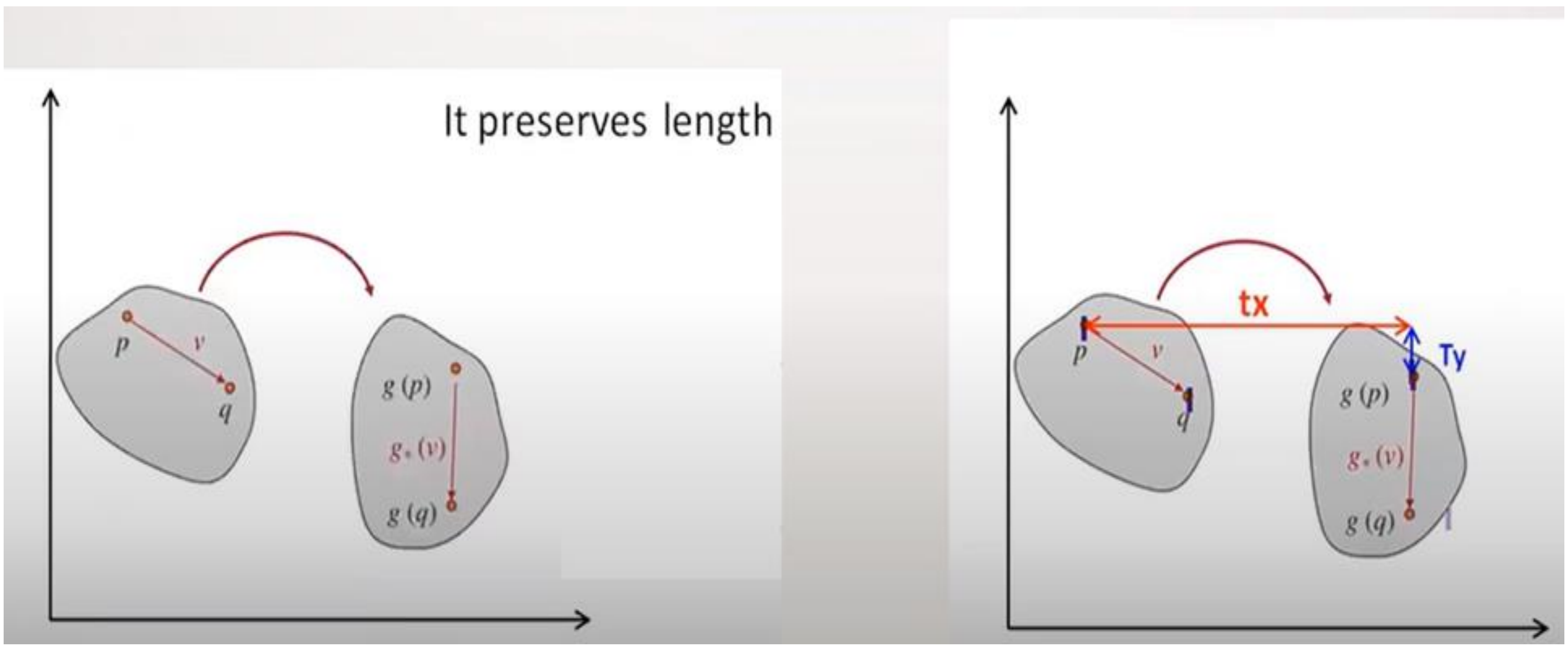
- Shearing an image means shifting the pixel values either horizontally or vertically.
- Basically, this shits some part of an image to one direction. Horizontal shearing will shift the upper part to the right and lower part to the left.
- Here you can see in gif. That upper part has shifted to the right and the lower part to the left.

Cont…



5.Rigid Transformation

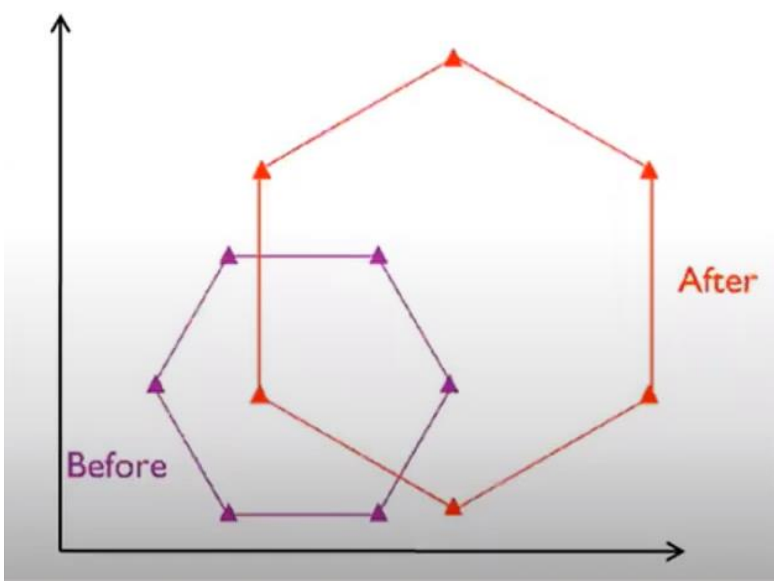
- Rigid = Translations + Rotations

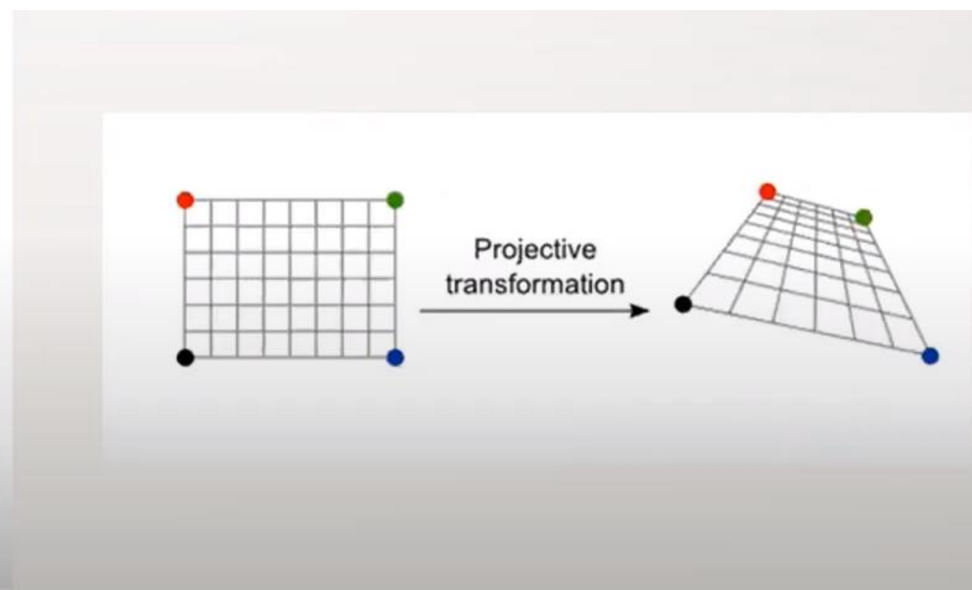


6. Similarity Transformation

- Similarity = Translations + Rotations + Scale

$$\begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = \begin{bmatrix} s\cos\theta & -s\sin\theta & t_x \\ s\sin\theta & s\cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \qquad \text{where } s = \text{scaling}$$



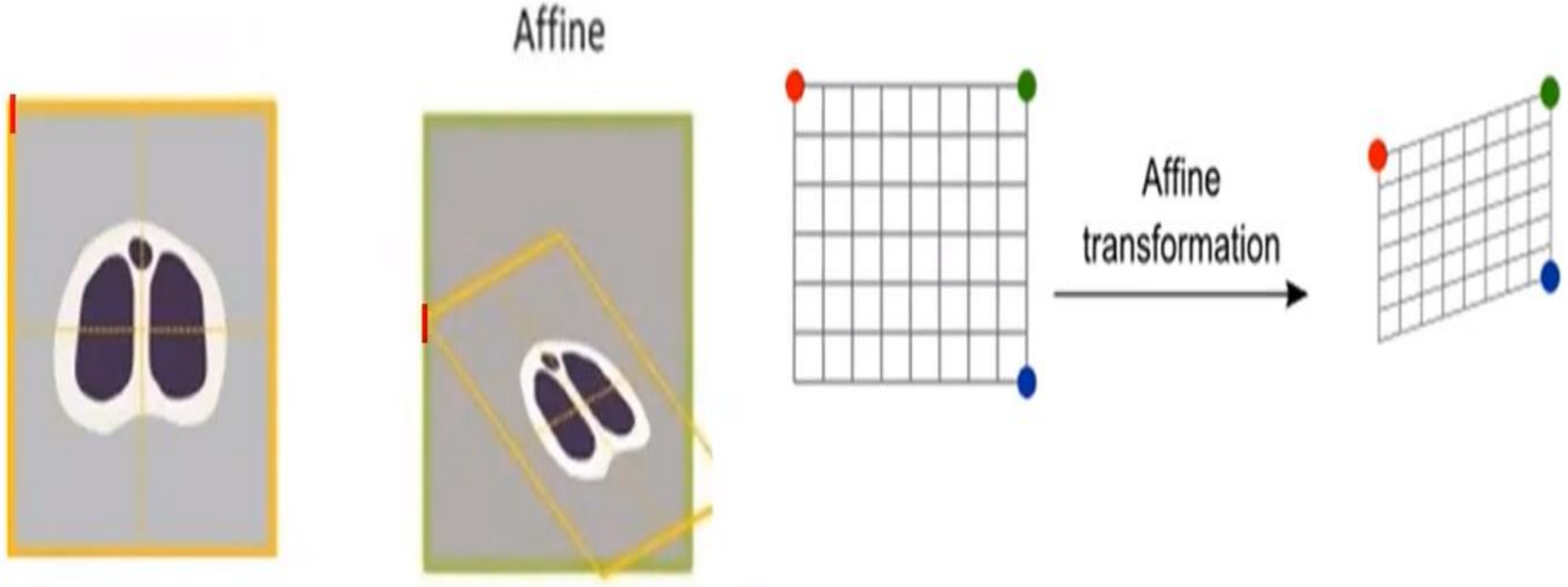


7.Affine Transformation

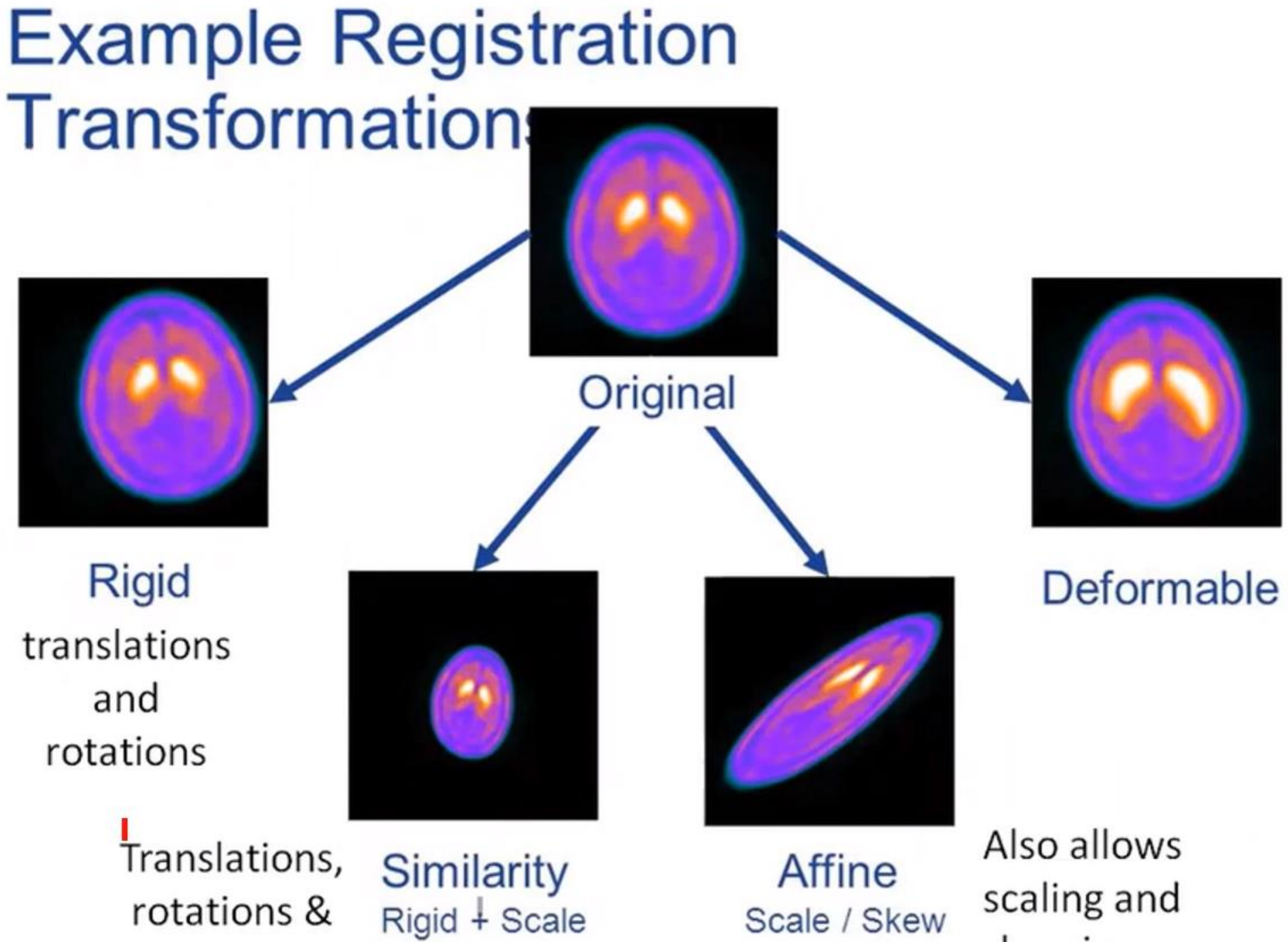
- Affine = Translations + Rotations + Scale + shear
- An affine transformation is a transformation that preserves co-linearity and the ration of distances.
- The parallel lines in an original image will be parallel in the output image.

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} a_0 & a_1 & a_2 \\ b_0 & b_1 & b_2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Cont…



Example Registration Transformations

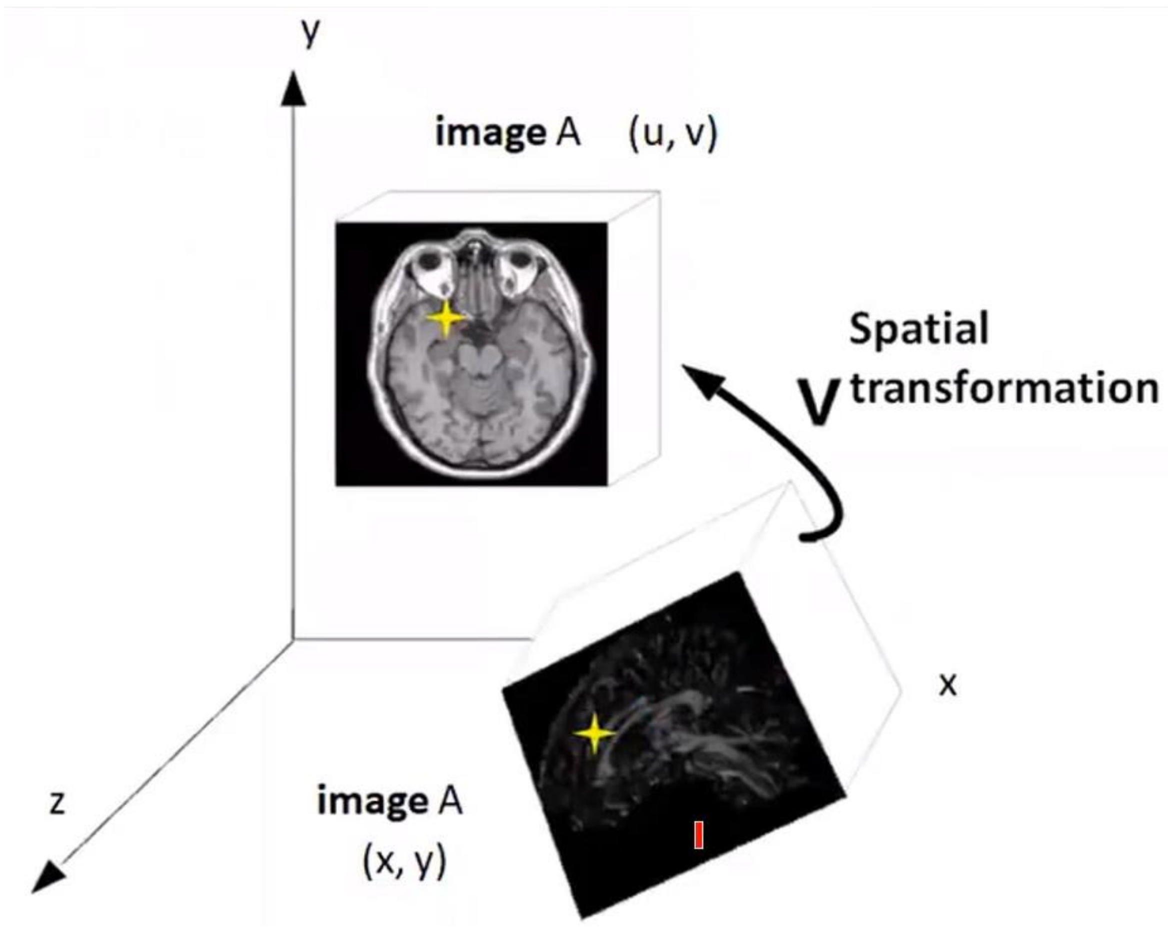


translations and rotations

Translations, rotations & scale

Also allows scaling and shearing

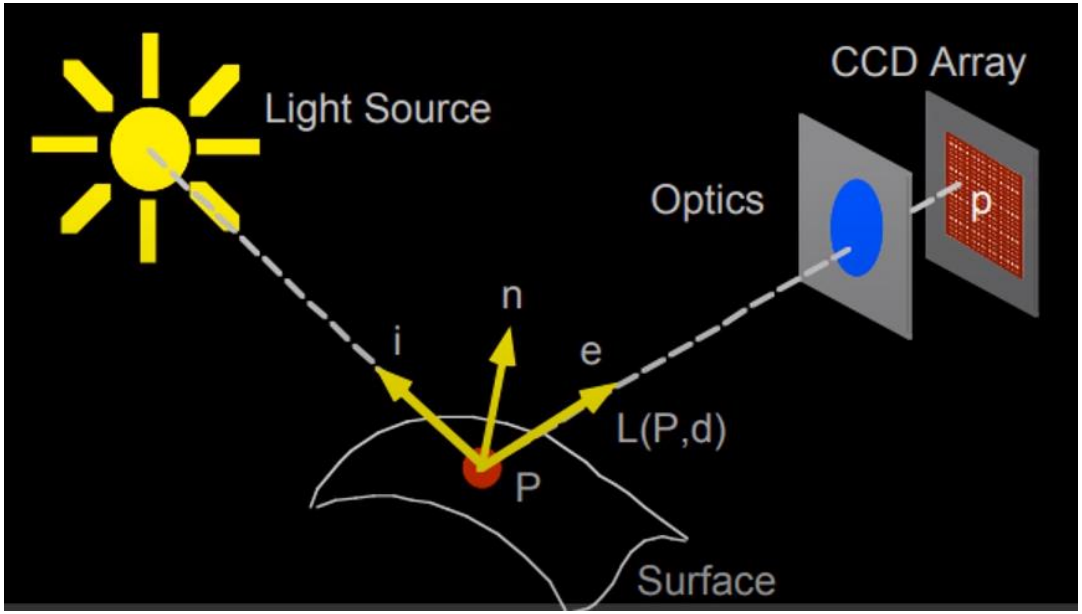
3D spatial Transformation



Radiometry

What is Radiometry?

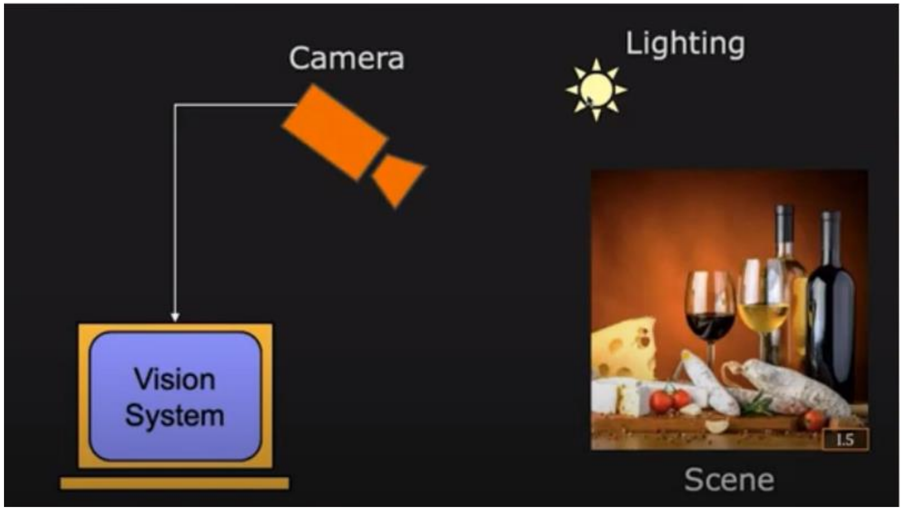
- Radiometry is the part of image formation concerned with the relation among the amounts of light energy emitted from light sources, reflected from surfaces, and registered by sensors.



Cont…

- Concerned with the relationship between the amount of light radiating from a surface and the amount incident at its image.
- In other words, what the brightness of the point will be.

From 3D to 2D


Camera
Lighting
Vision
System
Scene
1.5

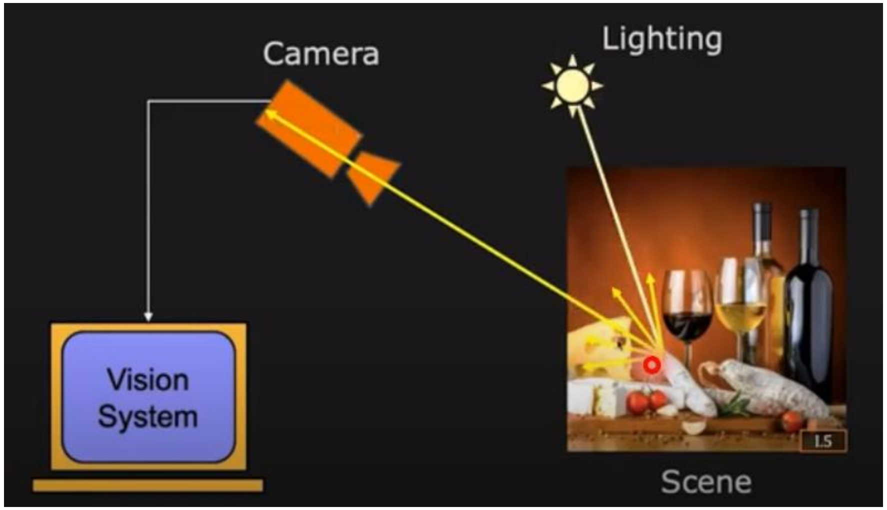
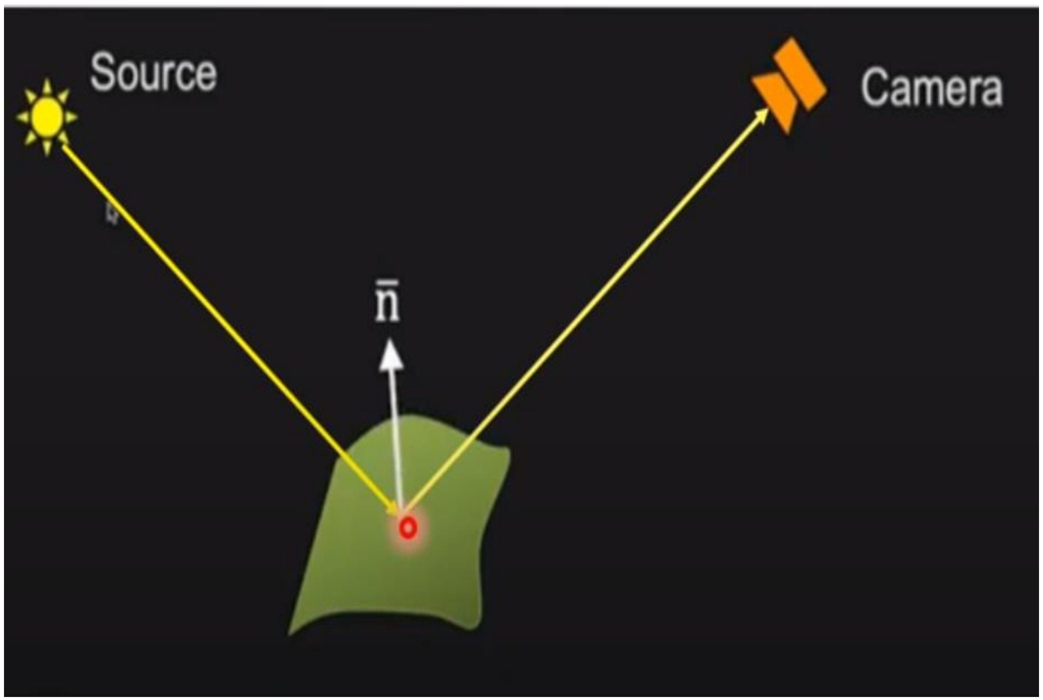
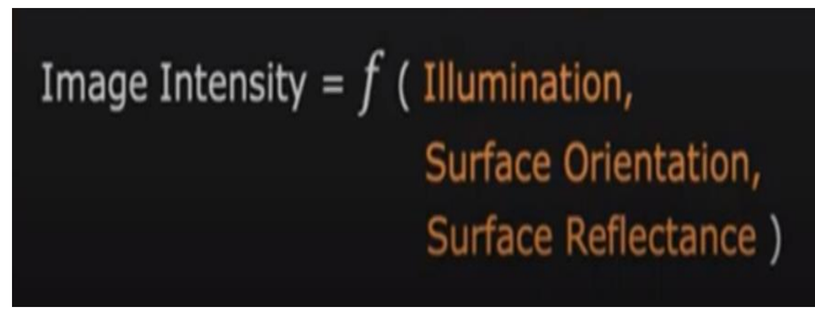

Camera
Lighting
Vision
System
Scene
1.5

Image Intensity

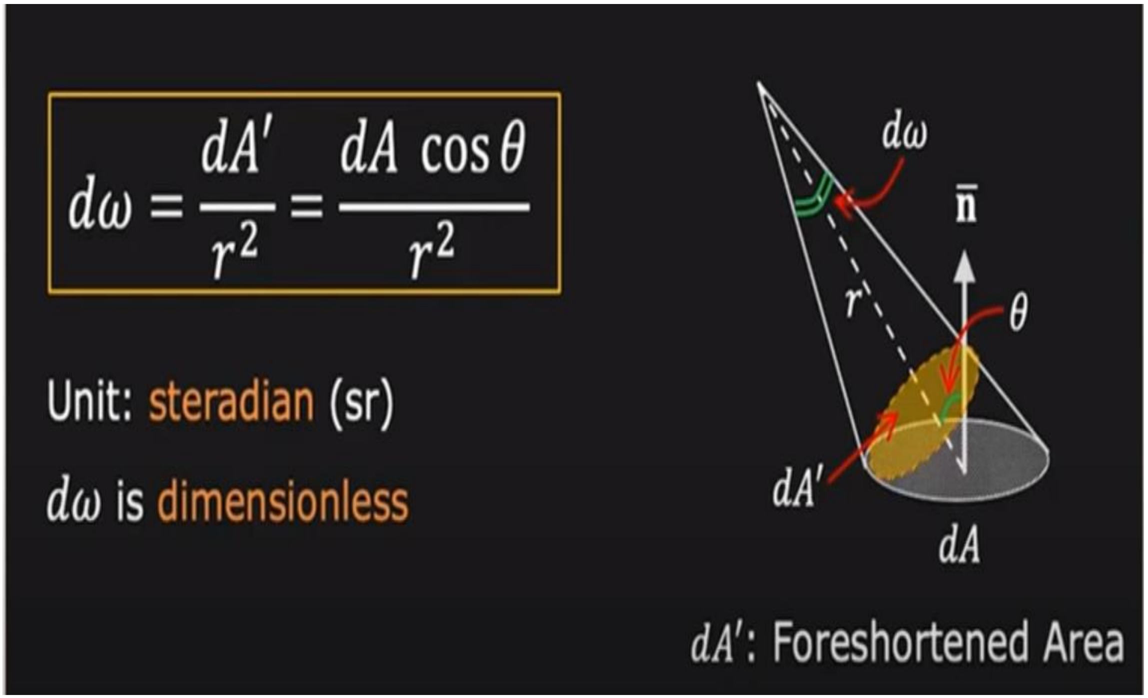
- Image intensity understanding is under-constrained.





Concept of Angle (2D):dθ

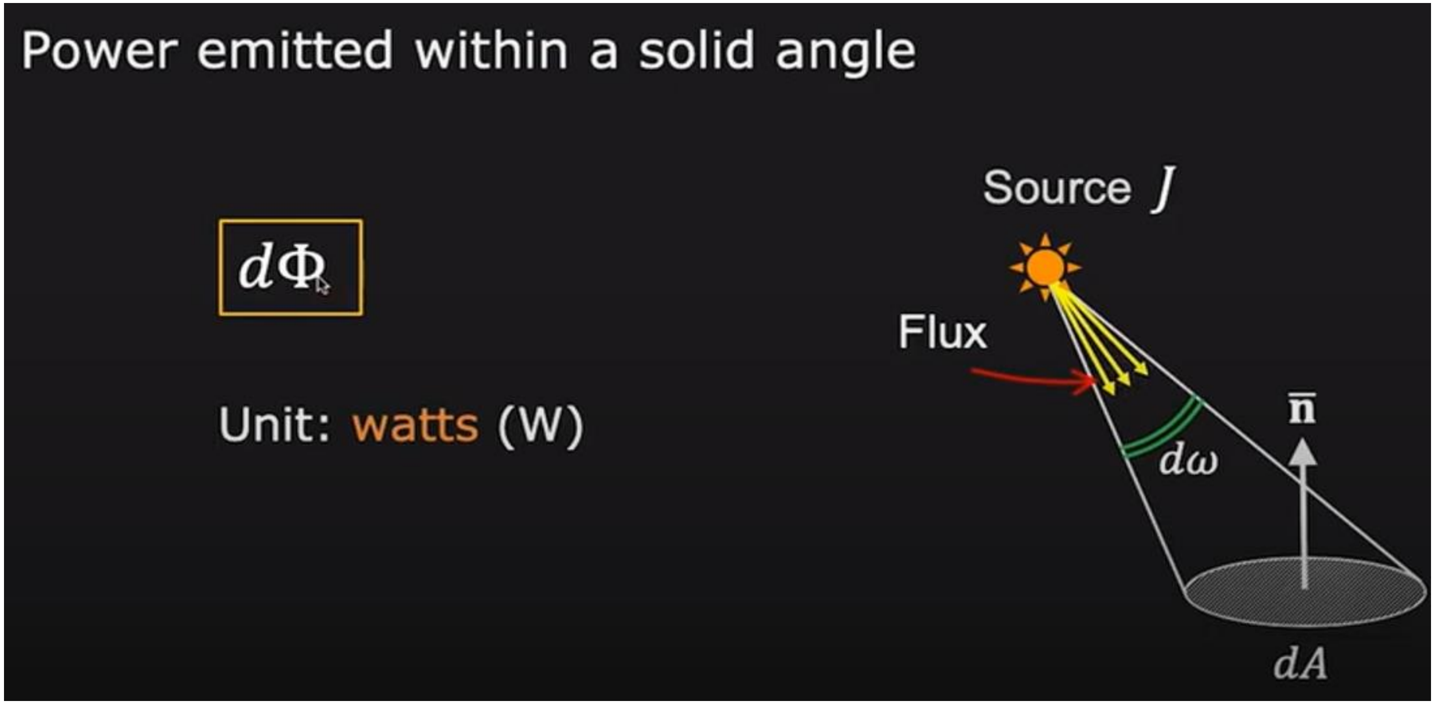
Solid Angle(3D):*dw*



1. Area of sphere $A = 4\pi r^2$
   - Solid angle of sphere = 4π
2. Area of hemisphere $A = 2\pi r^2$
   - Solid angle of sphere = 2π

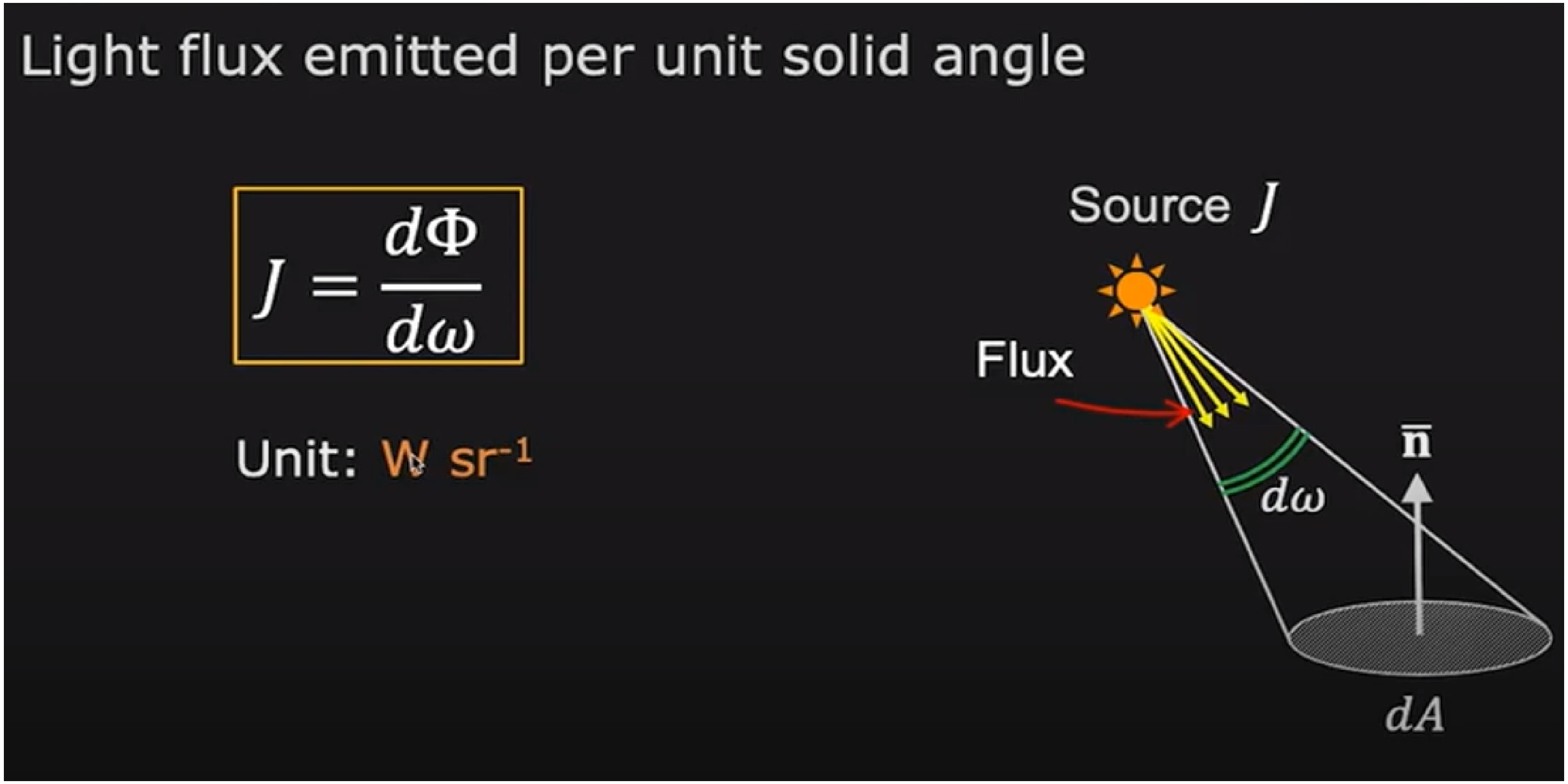
Light Flux :$d\phi$

- Luminous flux(in lumens) is a measure of the total amount of light a lamp puts out.



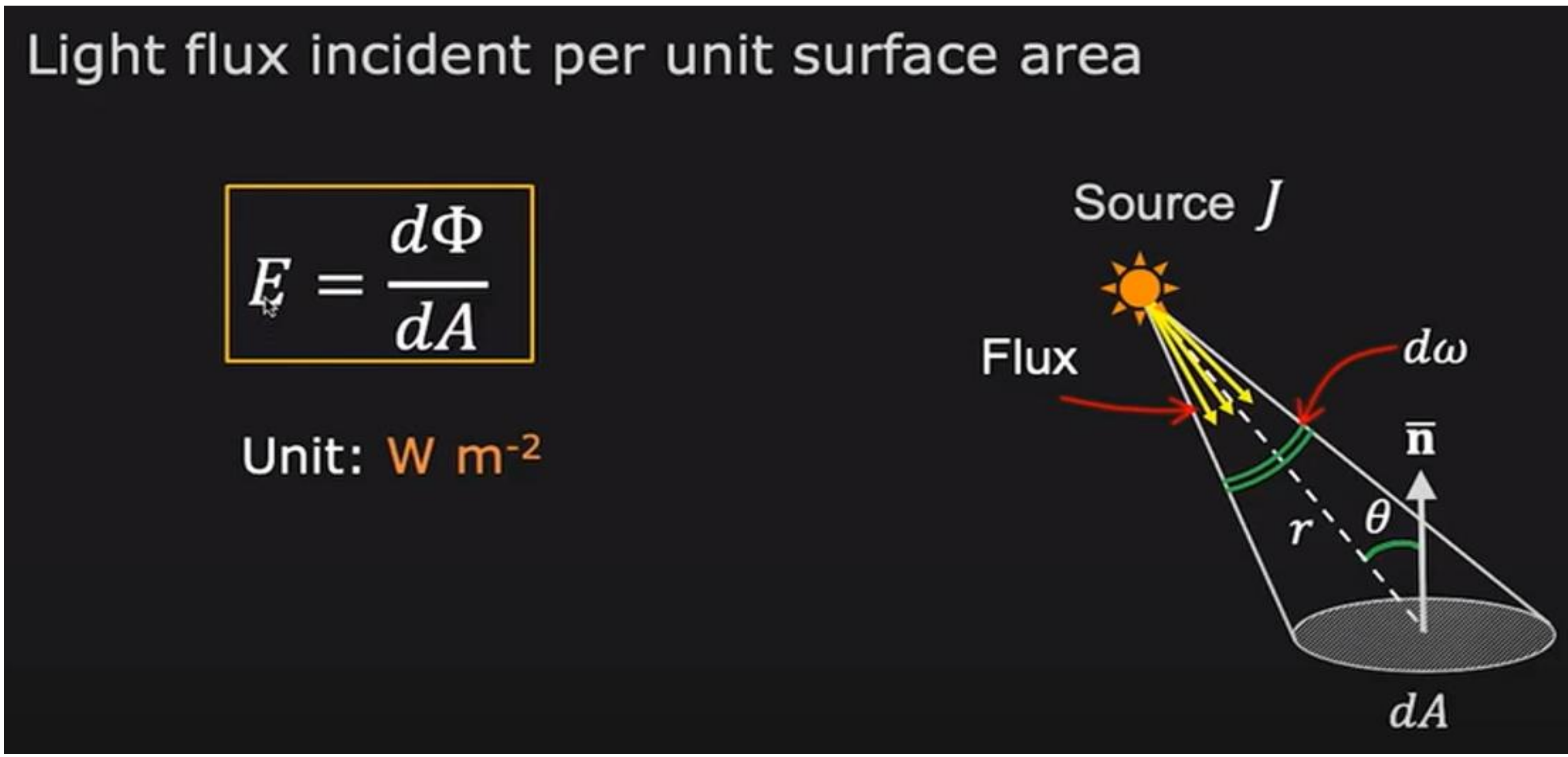
Radiant Intensity: *J*

- Brightness of source



Surface Irradiance: *E*

- Illumination of surface.



Cont…

$$E = \frac{d\Phi}{dA}$$

$$d\theta = \frac{dl}{r}$$

$$d\omega = \frac{dA'}{r^2} = \frac{dA\ \cos\theta}{r^2}$$

$$J = \frac{d\Phi}{d\omega}$$

$$E = \frac{d\Phi}{dA}$$

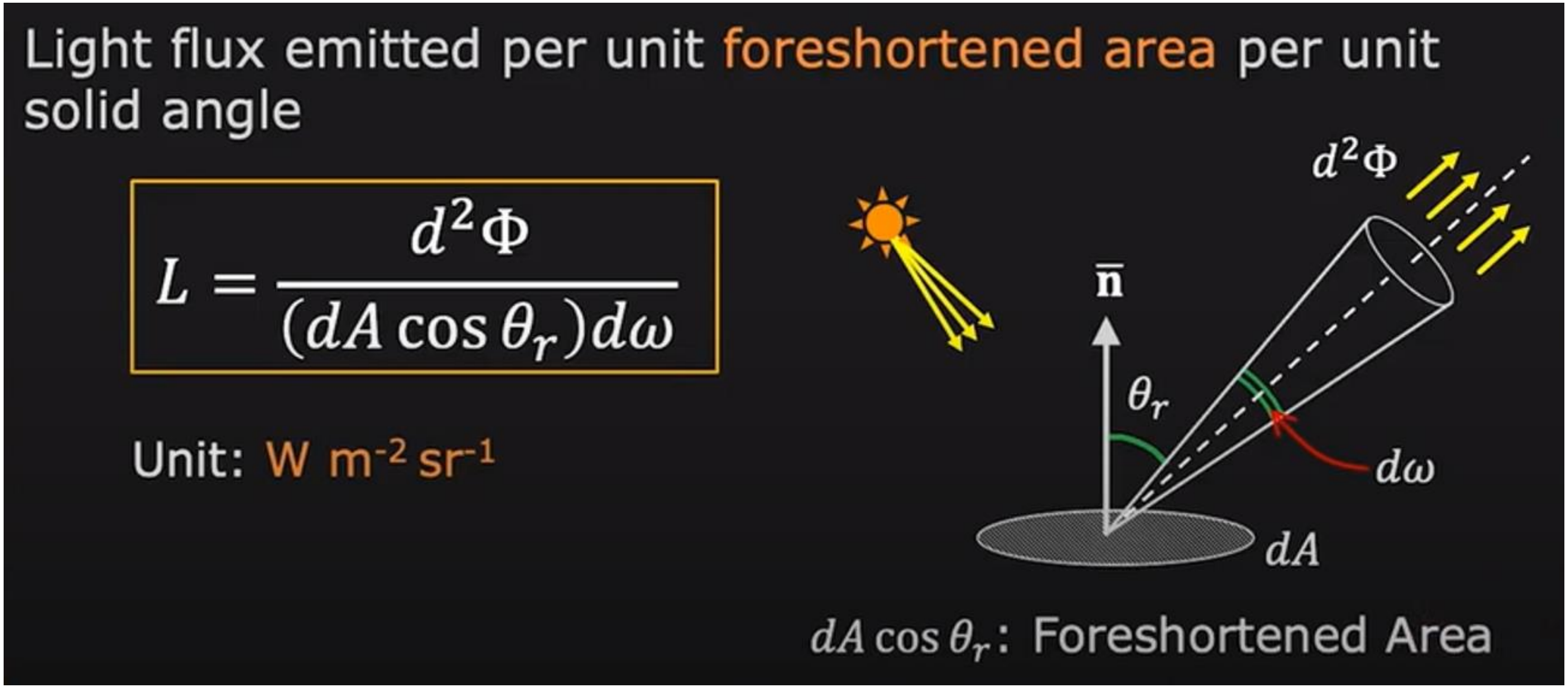
$$d\phi = J.d\omega$$

$$E = \frac{J.d\omega}{dA}$$

$$E = \frac{J\,d\omega}{dA} = \frac{J\,\frac{dA\cos\theta}{r^2}}{dA} = \frac{J\cos\theta}{r^2}$$

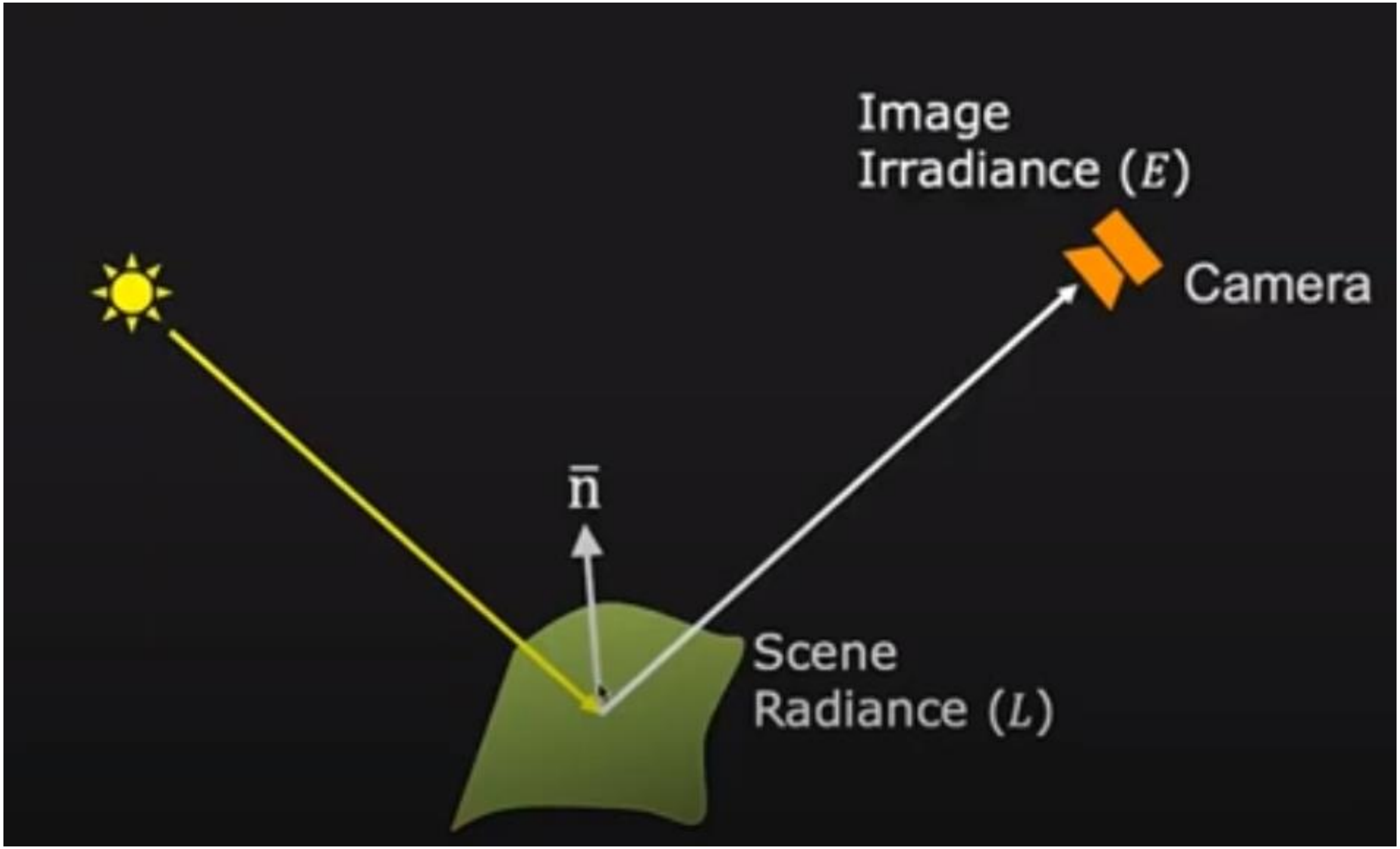
Surface Radiance : L

- Brightness of surface itself.

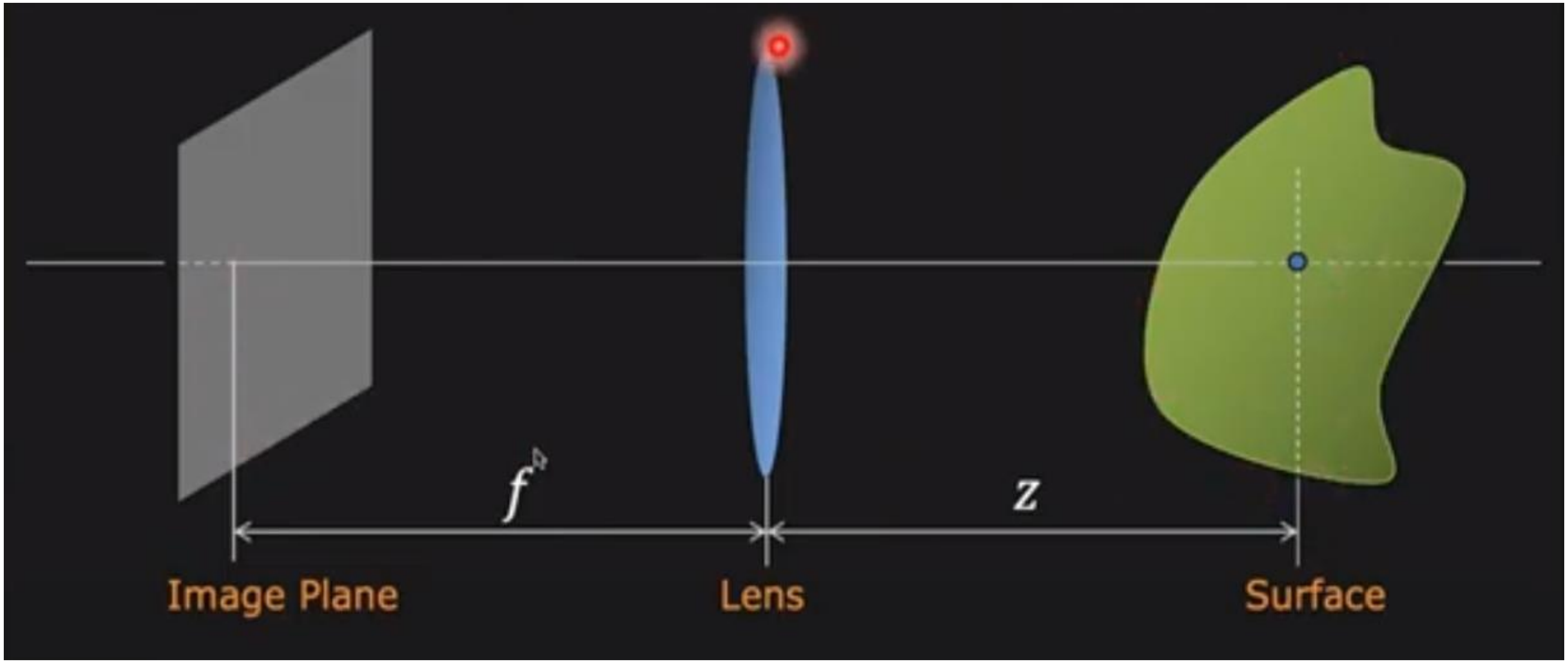


Radiometry: Scene Radiance and Image Irradiance

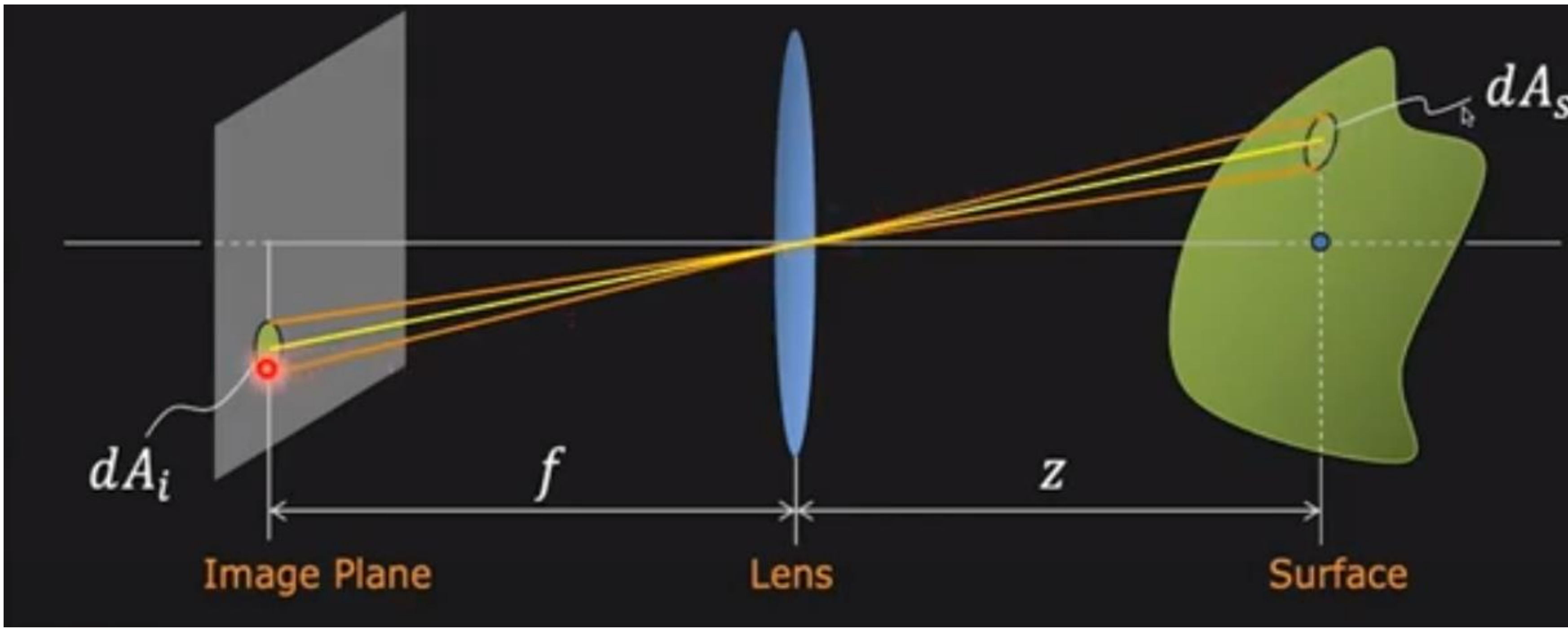
Scene Radiance and Image Irradiance



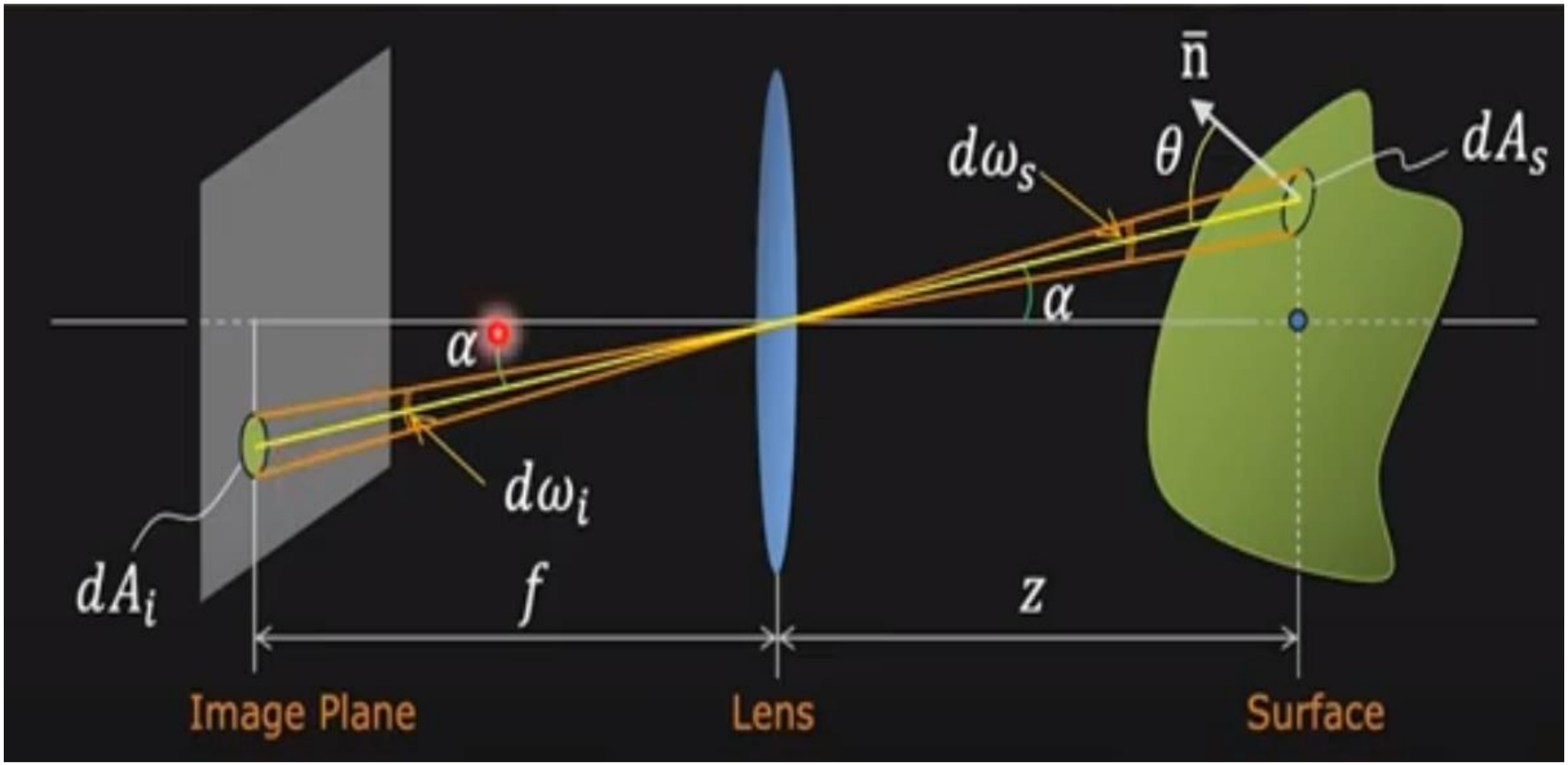
Cont…



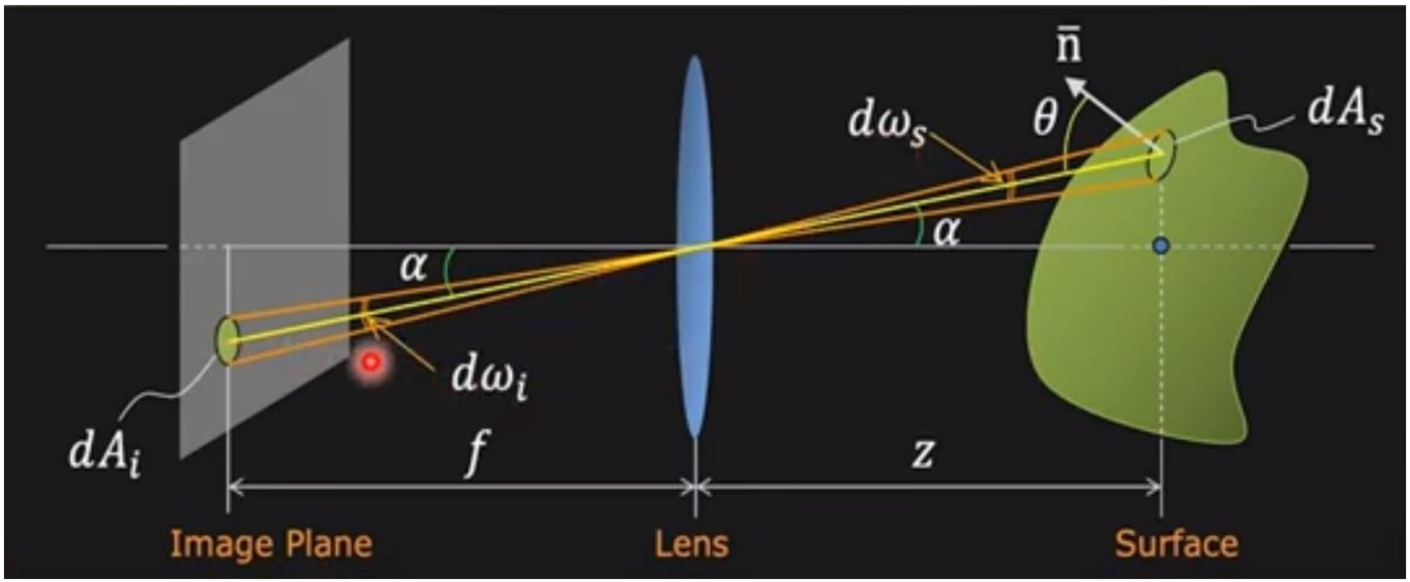
Cont…



Cont…



Cont…

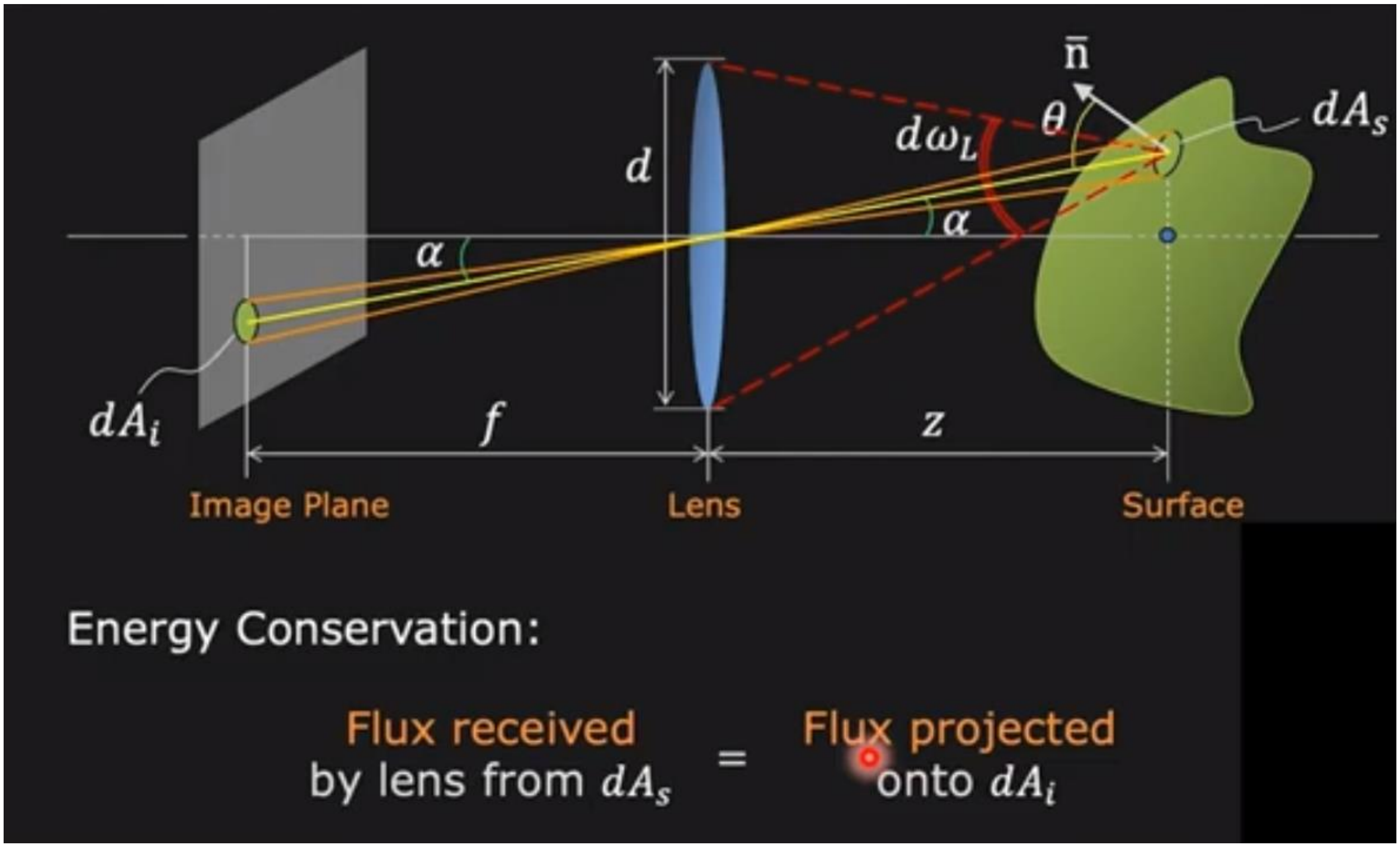


Solid Angles: $d\omega_i = d\omega_s$

$$\frac{dA_i \cos\alpha}{(f/\cos\alpha)^2} = \frac{dA_s \cos\theta}{(z/\cos\alpha)^2} \Rightarrow \boxed{\frac{dA_s}{dA_i} = \frac{\cos\alpha}{\cos\theta}\left(\frac{z}{f}\right)^2}$$

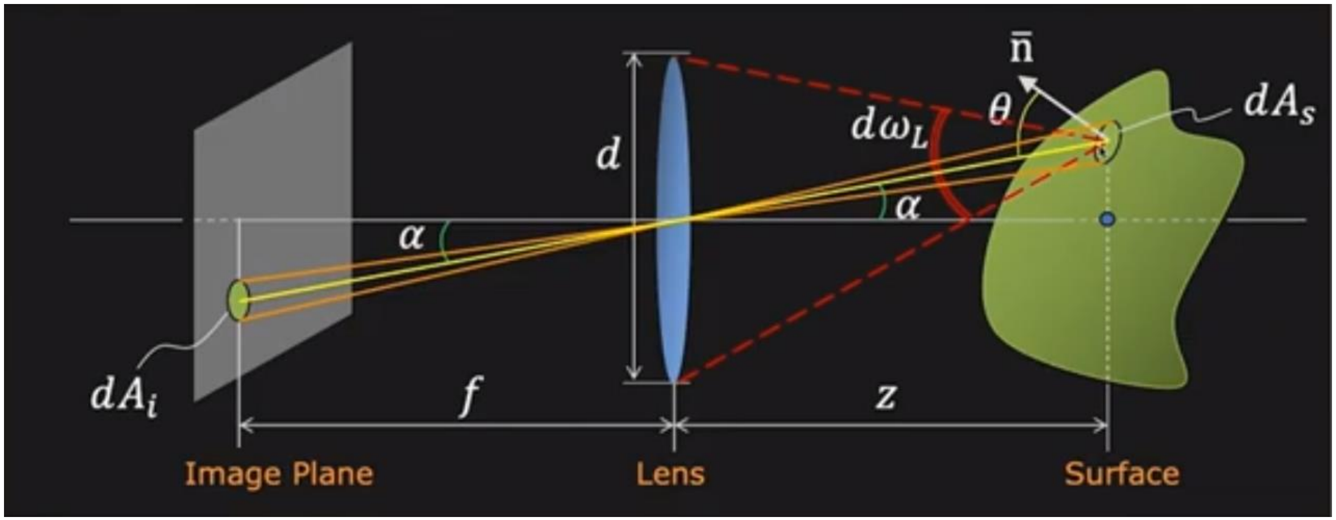
Equation (1)

Cont…



$\bar{n}$
dA_s
θ
$d\omega_L$
d
α
α
dA_i
f
z
Image Plane
Lens
Surface
Energy Conservation:
Flux received by lens from dA_s = Flux projected onto dA_i

Cont…



Scene Radiance:

$$L = \frac{d^2\Phi}{(dA_s \cos\theta)d\omega_L}$$

Flux received by lens from dA_s

$$\boxed{d^2\Phi = L\,(dA_s \cos\theta)d\omega_L}$$

Equation (3)

Cont…

Equation (1)

$$\frac{dA_s}{dA_i} = \frac{\cos\alpha}{\cos\theta}\left(\frac{z}{f}\right)^2$$

Equation (2)

$$d\omega_L = \frac{\pi d^2}{4}\frac{\cos\alpha}{(z/\cos\alpha)^2}$$

Equation (3)

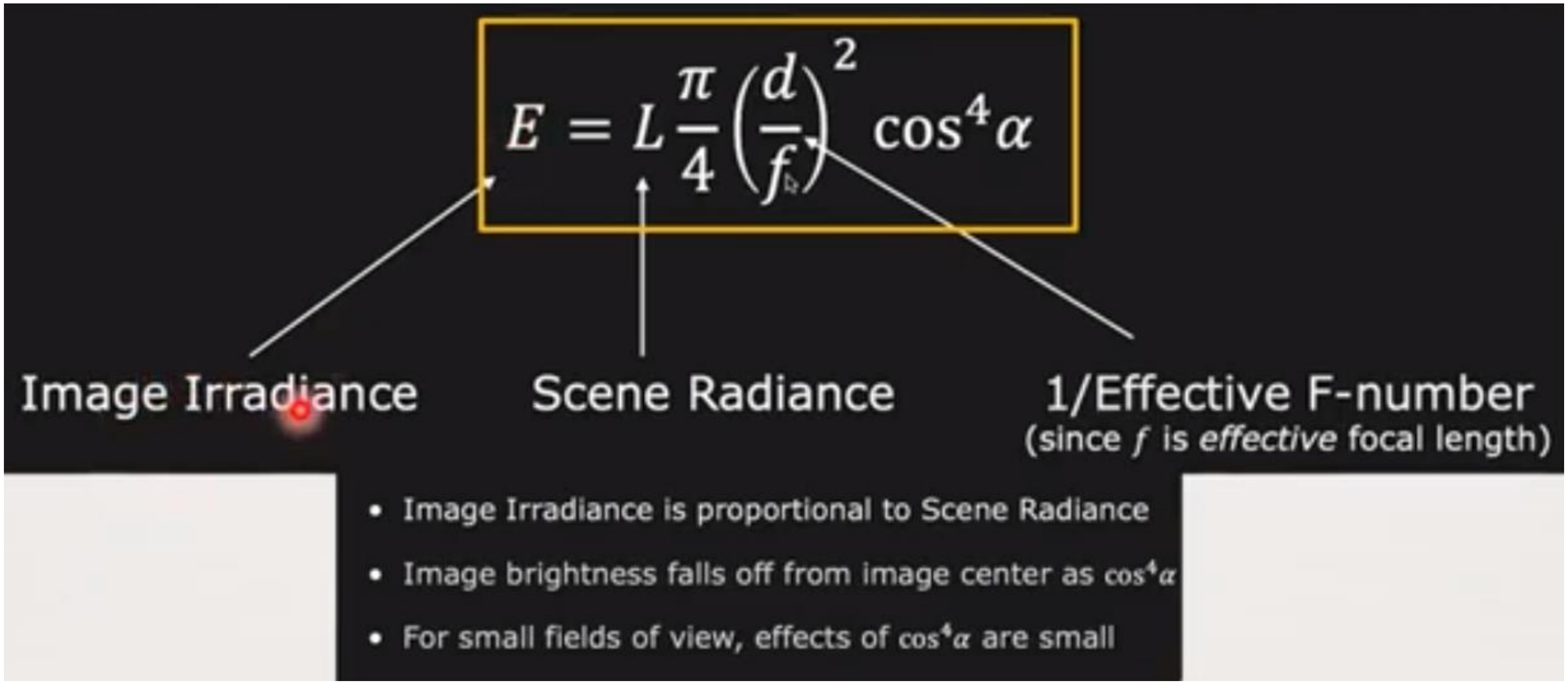
$$d^2\Phi = L\,(dA_s\cos\theta)\,d\omega_L$$

Equation (4)

$$d\Phi = E\,dA_i$$

$$\boxed{E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2\cos^4\alpha}$$

Cont…

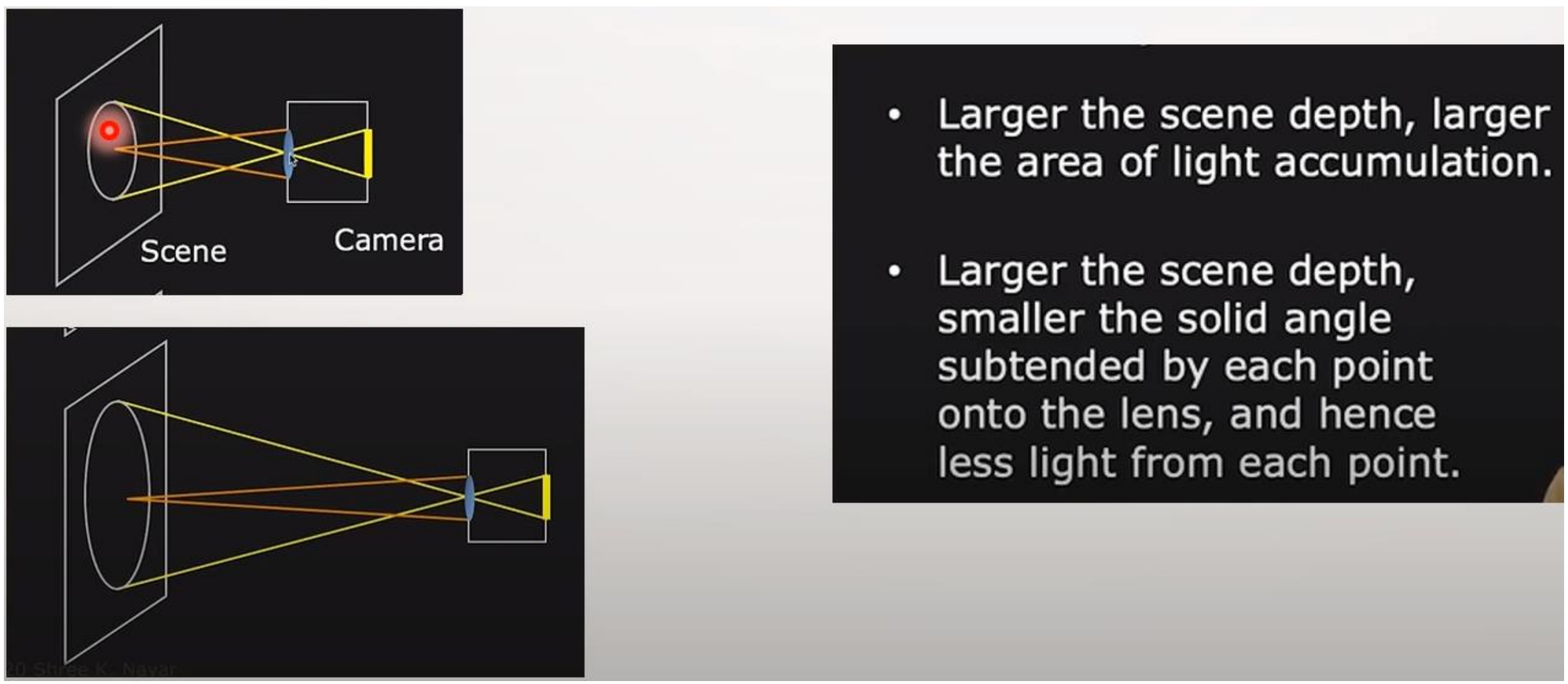


$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\alpha$
Image Irradiance
Scene Radiance
1/Effective F-number
(since f is *effective* focal length)
• Image Irradiance is proportional to Scene Radiance
• Image brightness falls off from image center as $\cos^4\alpha$
• For small fields of view, effects of $\cos^4\alpha$ are small

Cont…

$$E = L\frac{\pi}{4}\left(\frac{d}{f}\right)^2 \cos^4\alpha$$

- Does image brightness vary with scene depth? Nope



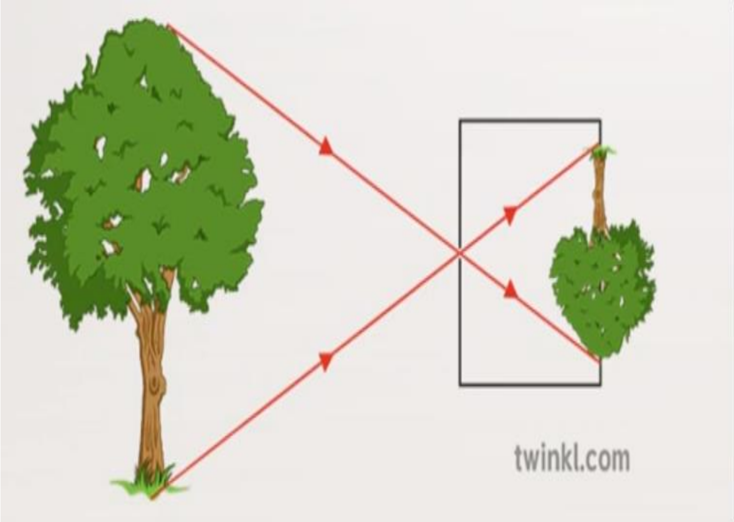
Cameras

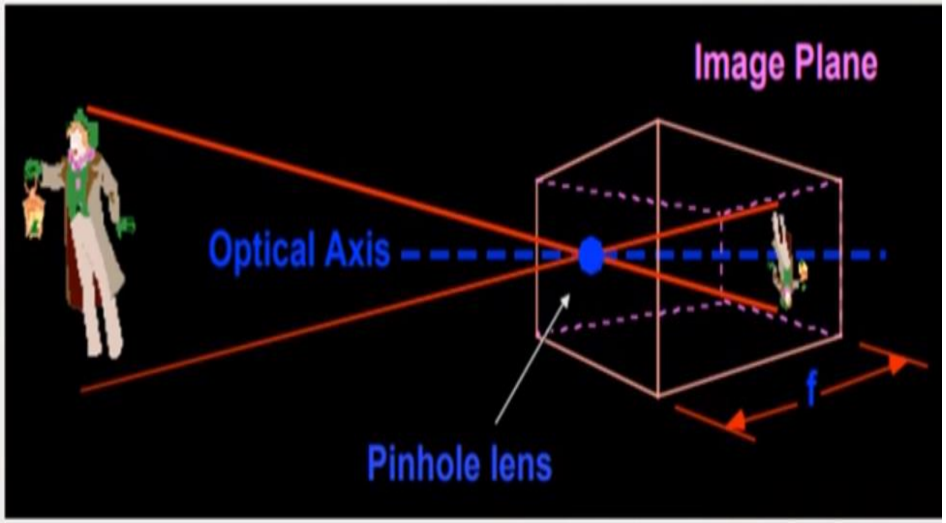
Cameras

- A camera is an optical instrument used to capture an image. At their most basic, cameras are sealed boxes(the camera body)with a small hole(the aperture)that light in to capture an image on a light-sensitive surface(usually photographic film or a digital sensor).
- Cameras have various mechanisms to control how the light falls onto the light sensitive surface. Lenses focus the light entering the camera, the size of the aperture can be widened or narrowed to let more or less light into the camera, and a shutter mechanism determines the amount of time the photo-sensitive surface is exposed to the light.

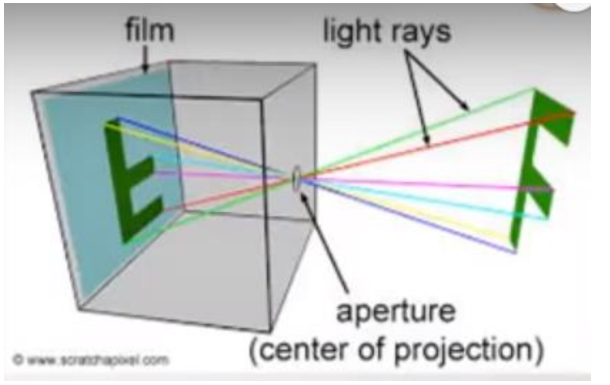
Pinhole Cameras Model

- A pinhole camera is "a simple camera without a lens and with a single small aperture." Many pinhole cameras are as a simple as a box with a hole in the side.

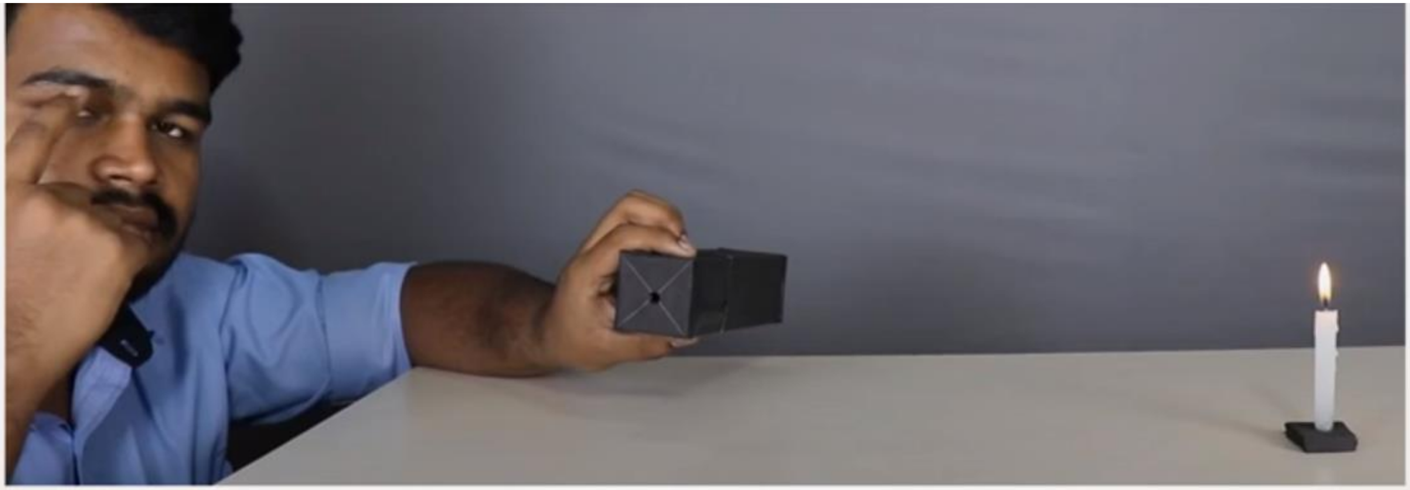


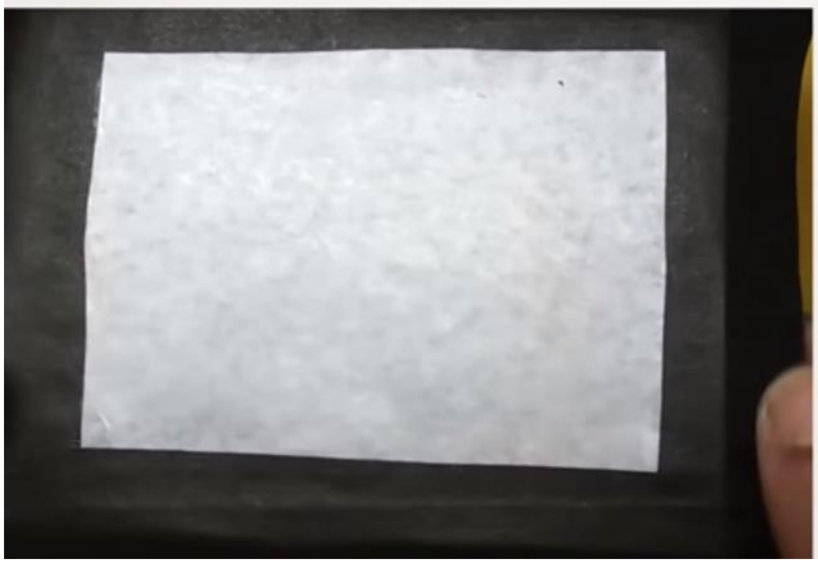


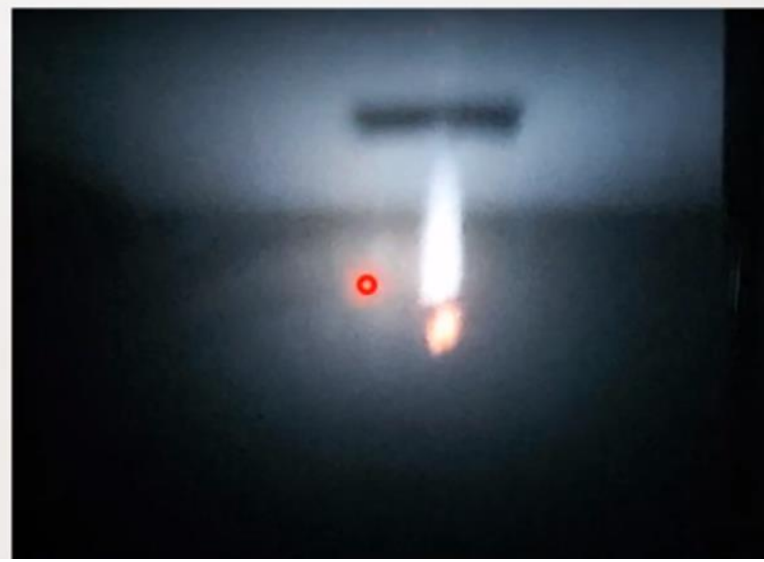
Pinhole Camera Model



- With a pinhole camera, this image is usually upside down and varies in clarity. Some people use a pinhole camera to study the movement of the sun over time(Solargraphy). A type of pinhole camera is often used to view an eclipse. Another type of pinhole camera, the camera obscure, was once used by artists.
- The device allowed the artist to view a scene through a different perspective. The artist would point the lens of the camera at the still life scene they wanted to paint. The camera would frame the image in smaller perspective thus allowing the artist to see the scene as it might appear. The person using a camera obscura might even trace the image on a piece of paper and achieve a very accurate copy of the scene.





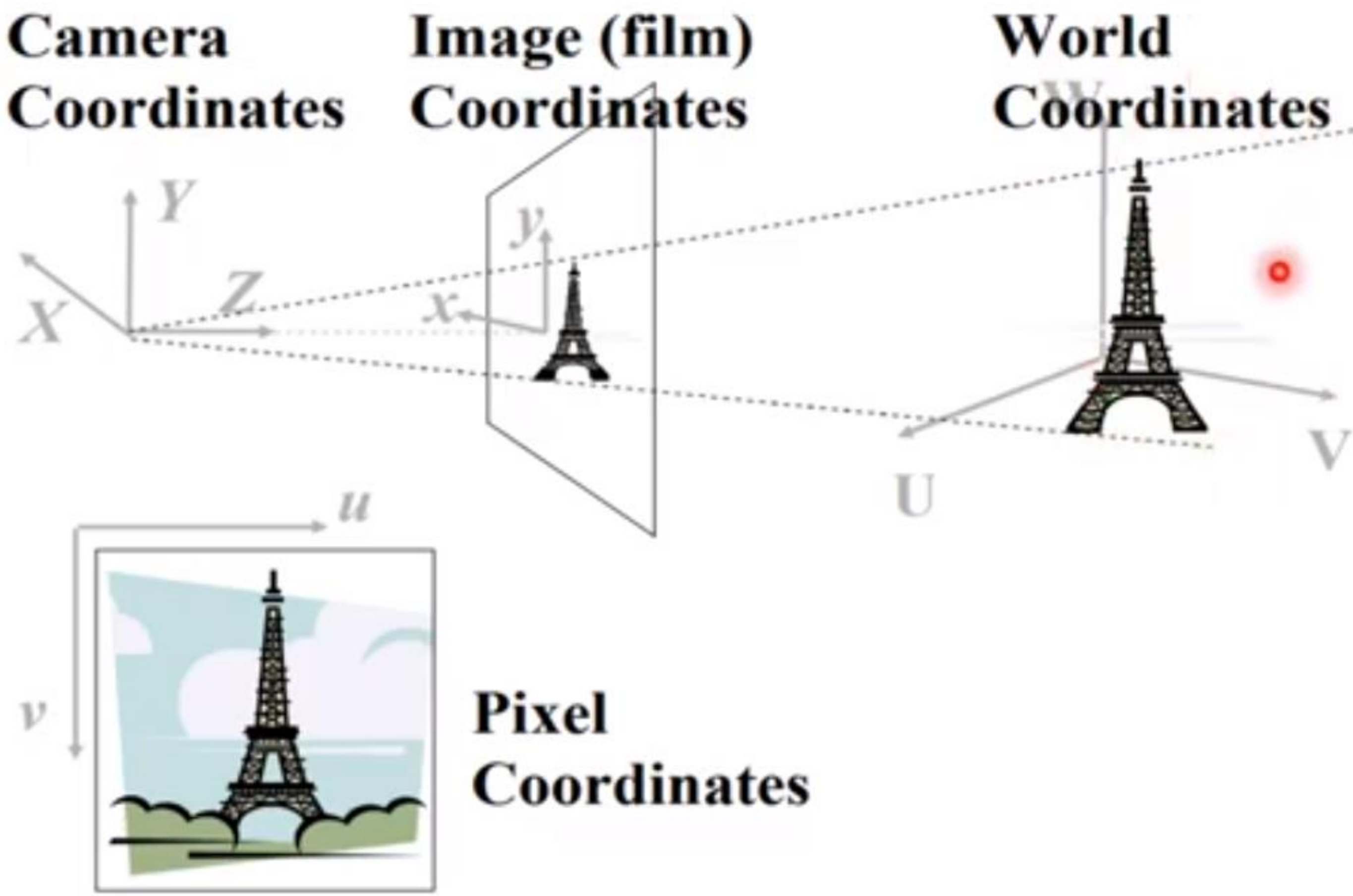


Projections

Projection of Image

Actual Image

Projections

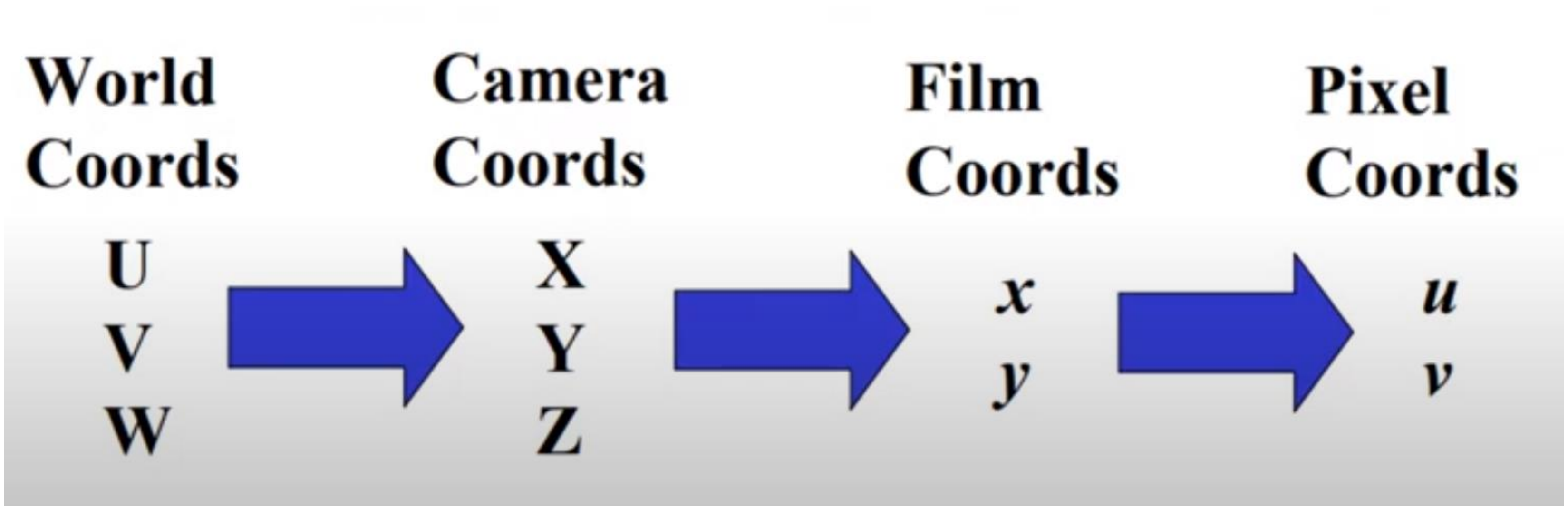


Cont…

1. Forward Projection
2. Backward Projection

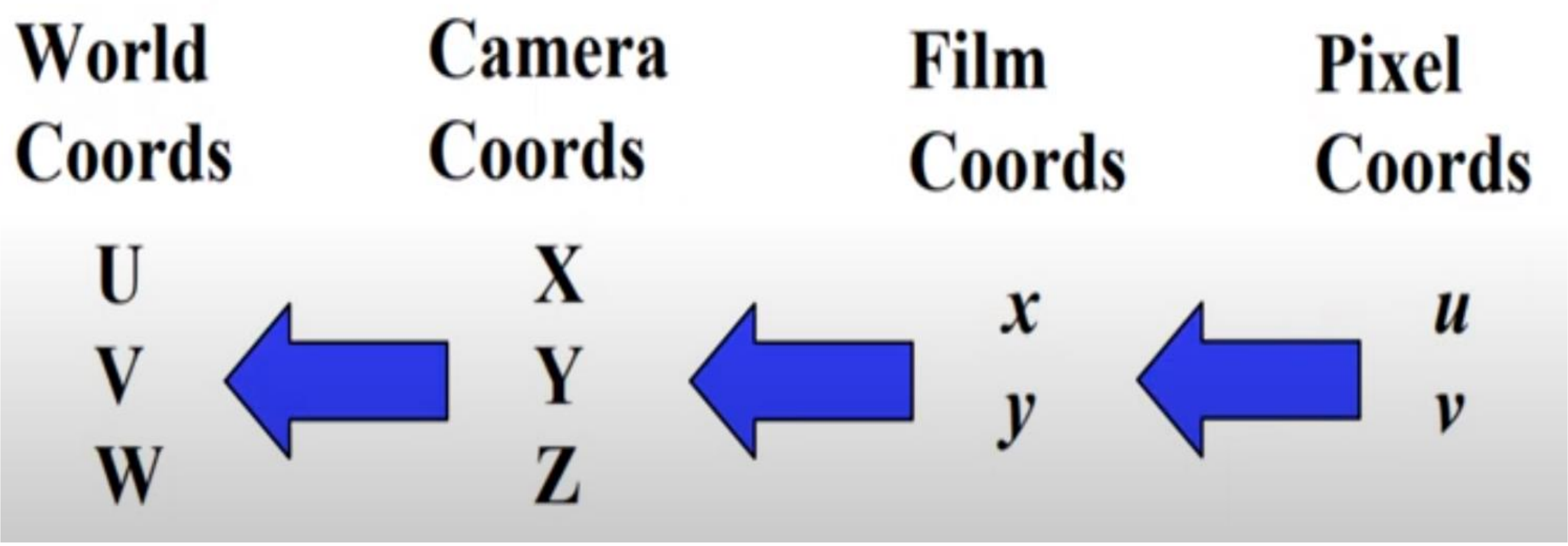
1. Forward Projection

- We want to mathematical model to describe how 3D world points get projected into 2D pixel coordinates.

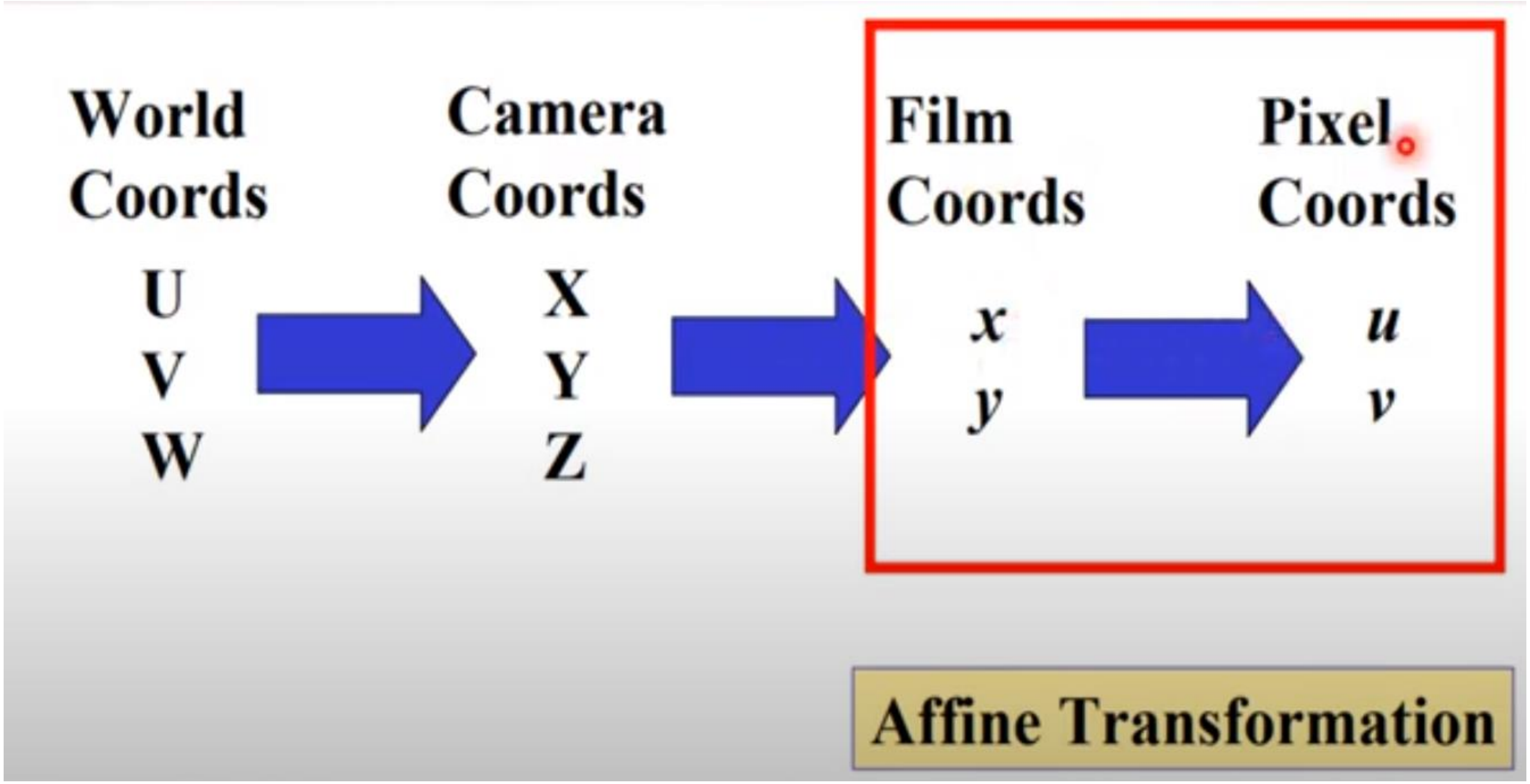


2. Backward Projection

- Recover 3D scene structure from image(via stereo or motion)



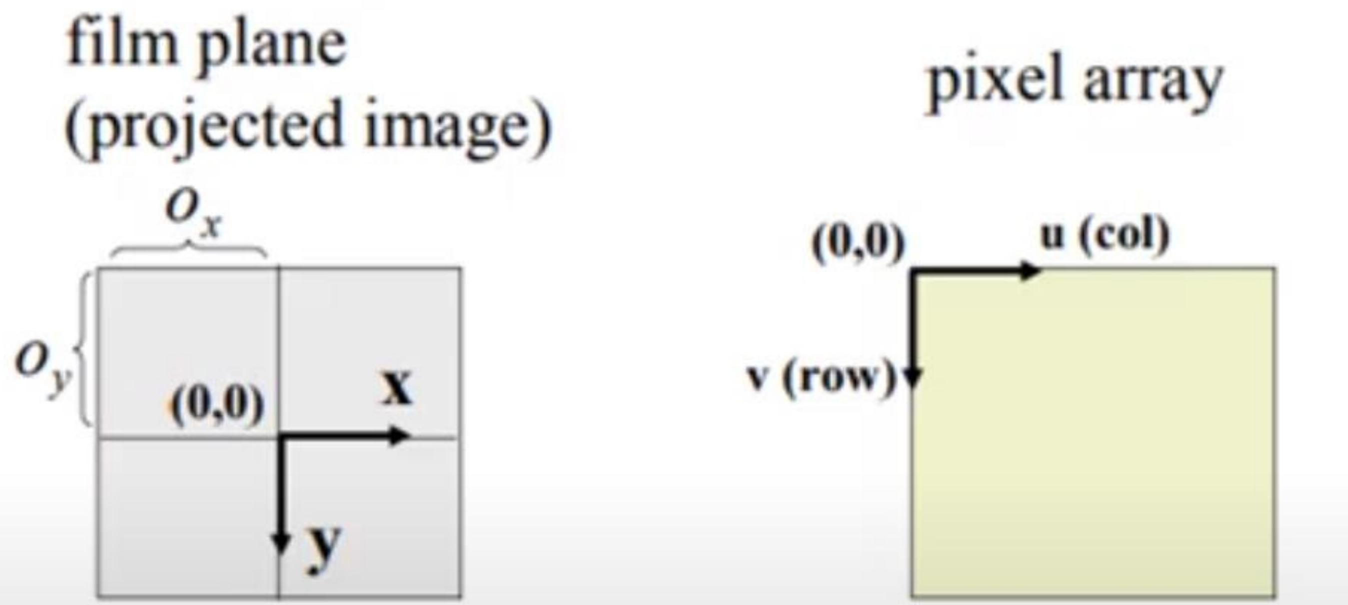
Intrinsic Camera Parameters



Cont…

- Describes coordinate transformation between film coordinates(projected image) and pixel array.
- Film cameras: scanning/digitization
- CCD cameras: grid of photo sensors

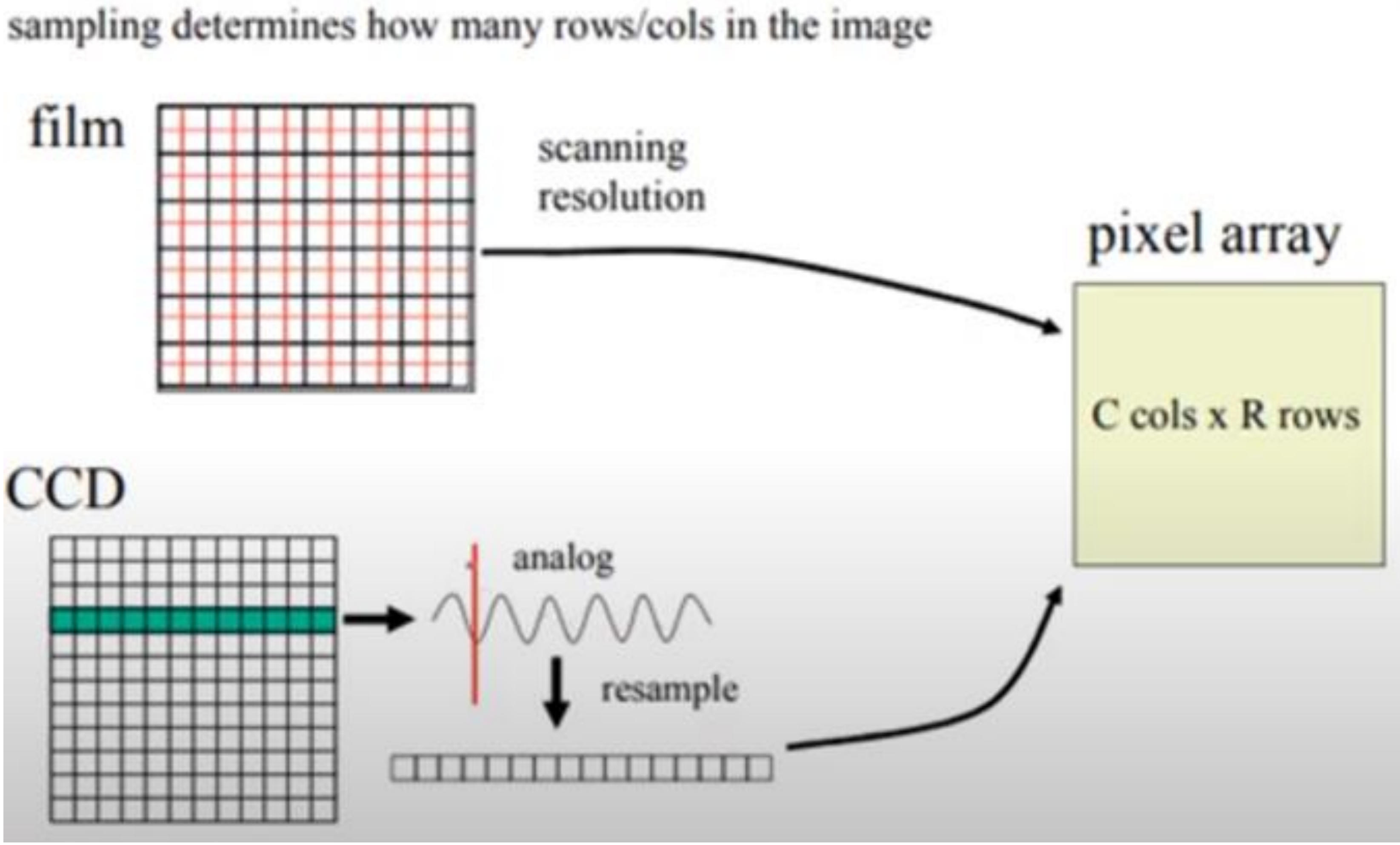
Intrinsic parameters(offsets)



$$u = f\frac{X}{Z} + o_x \qquad v = f\frac{Y}{Z} + o_y$$

o_x and o_y called image center or principle point

Intrinsic parameters(Scales)



Effective Scales: Sx and Sy

$$u = \frac{1}{s_x} f \frac{X}{Z} + o_x \qquad v = \frac{1}{s_y} f \frac{Y}{Z} + o_y$$

Note, since we have different scale factors in x and y, we don't necessarily have square pixels!

Aspect ratio is s_y / s_x

Thank you!!!