Chapter 1
Overview of computer vision and its applications

Prof. Mittal Darji
IT, GCET

Goal

- To create 'digital' images which can be processed to recover some of the characteristics of the 3D world which was imaged.

Assumptions

Typical imaging scenario:

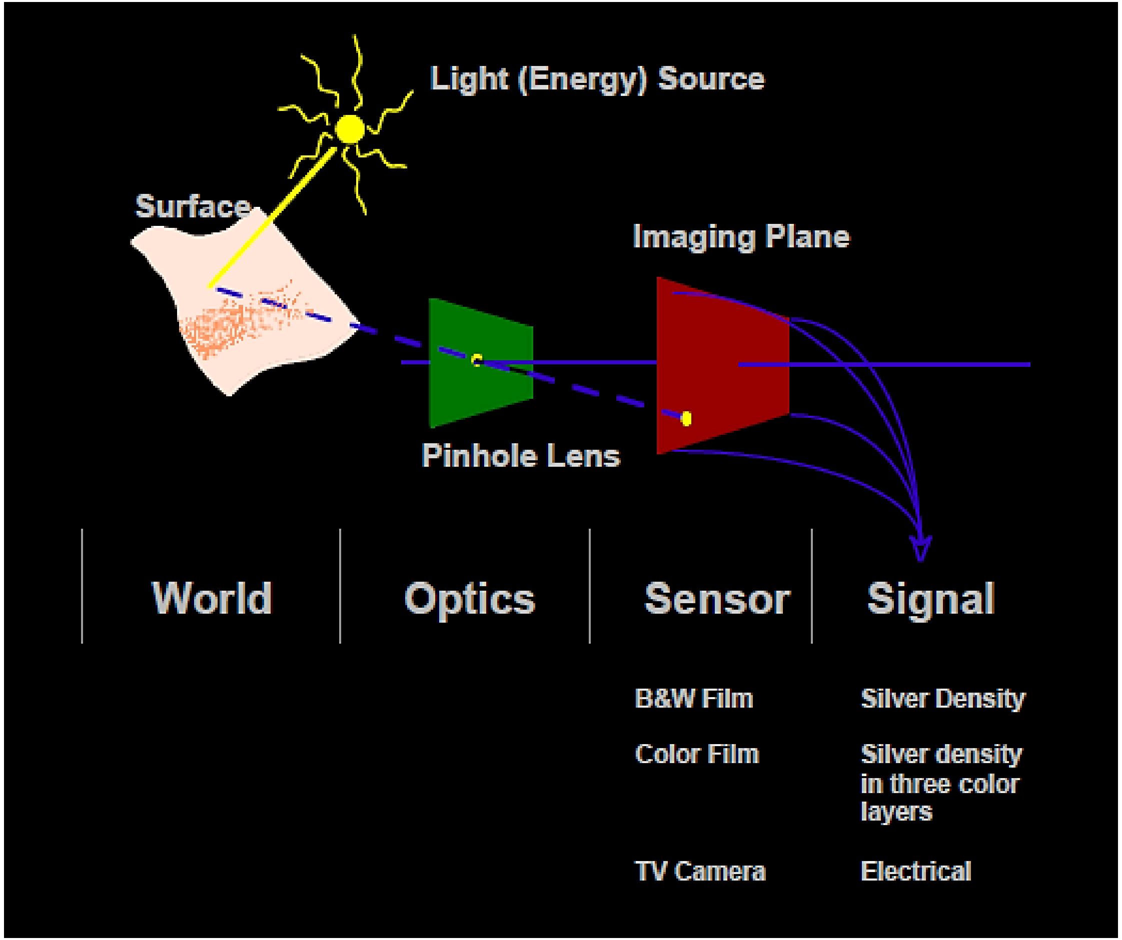
- visible light
- ideal lenses
- standard sensor (e.g. TV camera)
- opaque objects

Steps of Image Formation

World → Optics → Sensor → Signal → Digitizer → Digital Representation

World	reality
Optics	focus light from world on sensor
Sensor	converts light to electrical energy
Signal	representation of incident light as continuous electrical energy
Digitizer	converts continuous signal to discrete signal
Digital Rep.	final representation of reality in computer memory

Image Formation



Factors in Image Formation

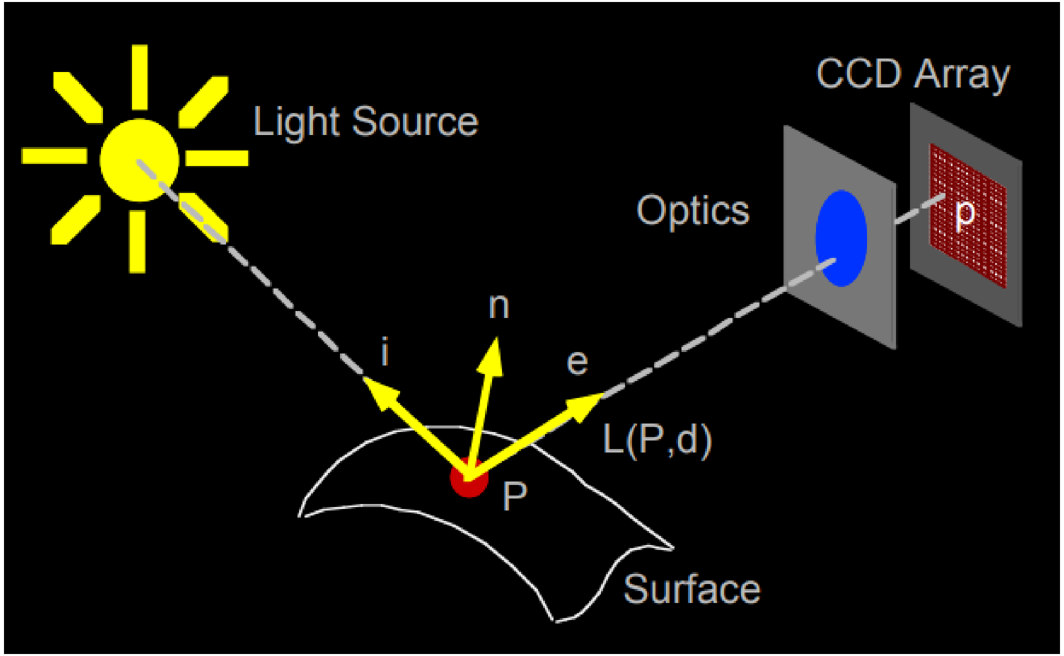
- **Geometry** - concerned with the relationship between points in the three-dimensional world and their images
- **Radiometry** - concerned with the relationship between the amount of light radiating from a surface and the amount incident at its image
- **Photometry** - concerned with ways of measuring the intensity of light
- **Digitization** - concerned with ways of converting continuous signals (in both space and time) to digital approximations

Geometry

- Geometry describes the projection of three-dimensional (3D) world → two-dimensional (2D) image plane
- Various kind of projections:
  - Perspective
  - Orthographic
  - Oblique
  - Isometric
  - Spherical

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.



Terminologies

- Brightness: informal notion used to describe both scene and image brightness.
- Image brightness: related to energy flux incident on the image plane:

  IRRADIANCE (illuminance)

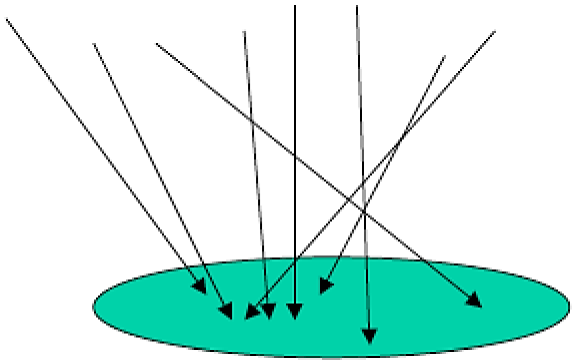
  "It's a bright day."
- Scene brightness: brightness related to energy flux emitted (radiated) from a surface.

  RADIANCE (luminance)

  "Yikes, that shirt is way too bright!"

Irradiance

- Irradiance is the power per unit area (Watts per square meter) of radiant energy falling on a surface.



$$E = \frac{dP}{dA}$$

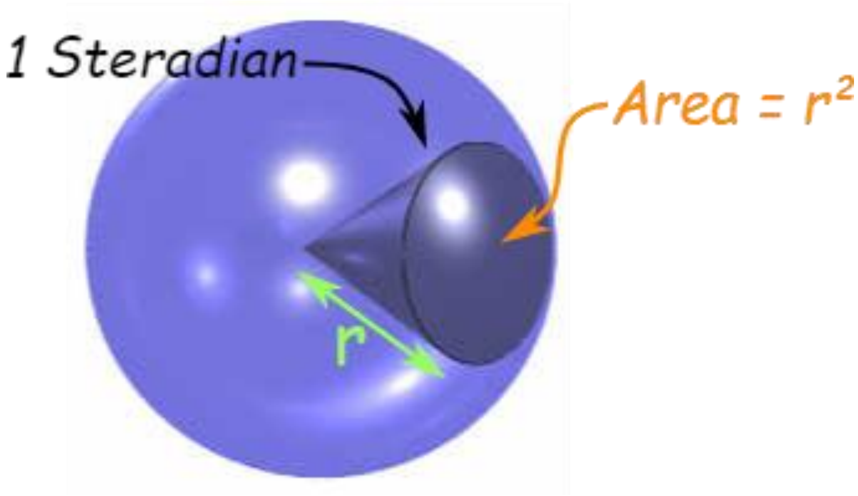
Irradiance

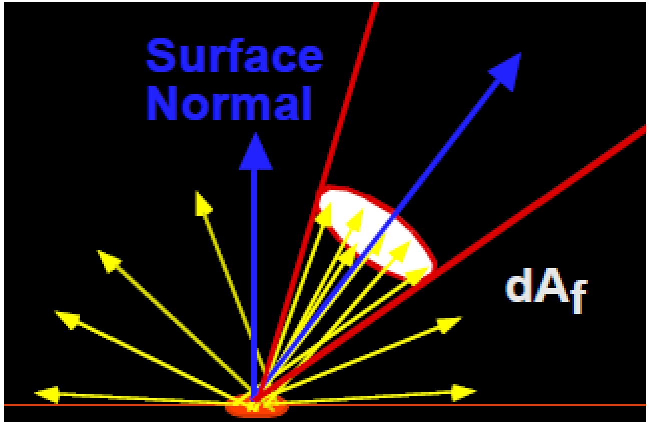
Radiance

- Radiance is the power emitted per unit area into a cone of directions having unit solid angle

(Watts per square meter per steradian.)

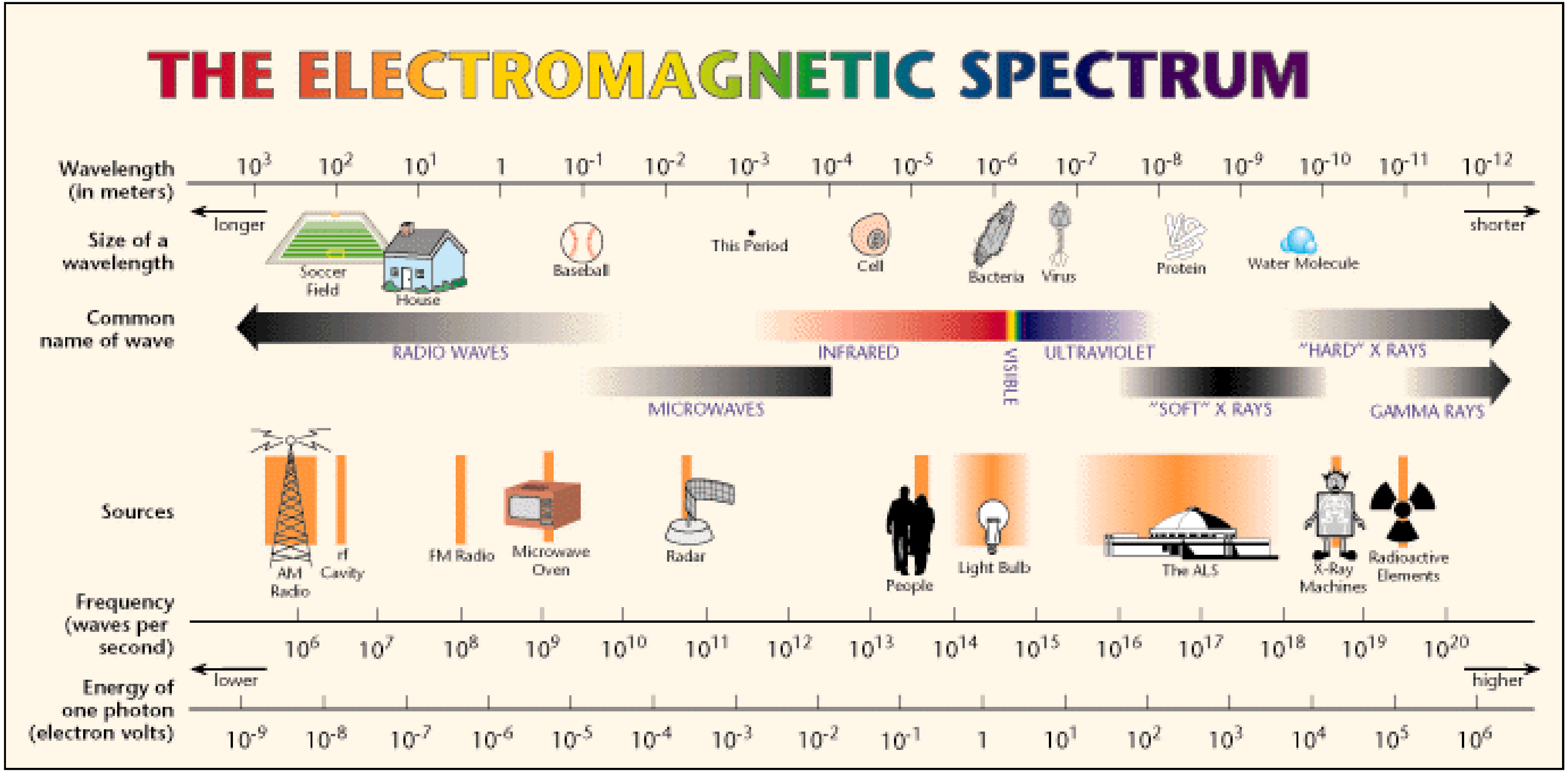
- 4pi steradians in a sphere
- Area of a unit sphere is 4pi steradians.
- How many steradians in a hemisphere?
- 4pi * 0.5 = 2pi steradians





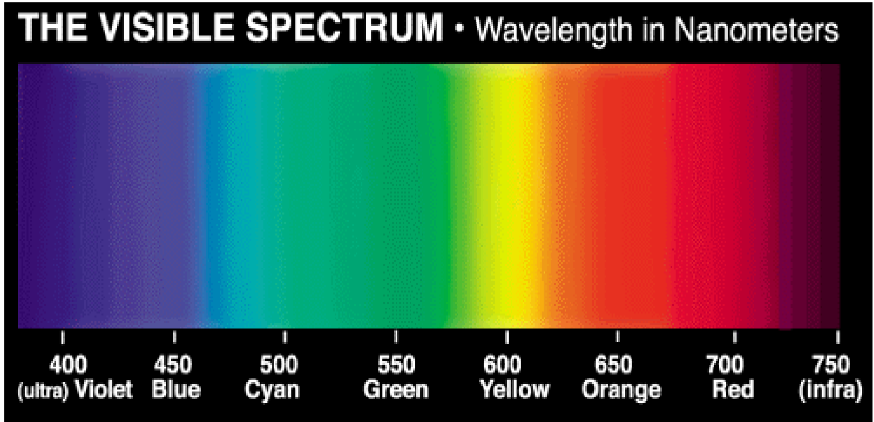
Light sources

- Point source
- Extended source
- Single wavelength
- Multi-wavelength
- Uniform
- Non-uniform


THE ELECTROMAGNETIC SPECTRUM
Wavelength (in meters)
10^3 10^2 10^1 1 10^-1 10^-2 10^-3 10^-4 10^-5 10^-6 10^-7 10^-8 10^-9 10^-10 10^-11 10^-12
longer
shorter
Size of a wavelength
Soccer Field
House
Baseball
This Period
Cell
Bacteria
Virus
Protein
Water Molecule
Common name of wave
RADIO WAVES
INFRARED
VISIBLE
ULTRAVIOLET
"HARD" X RAYS
MICROWAVES
"SOFT" X RAYS
GAMMA RAYS
Sources
AM Radio
rf Cavity
FM Radio
Microwave Oven
Radar
People
Light Bulb
The ALS
X-Ray Machines
Radioactive Elements
Frequency (waves per second)
10^6 10^7 10^8 10^9 10^10 10^11 10^12 10^13 10^14 10^15 10^16 10^17 10^18 10^19 10^20
lower
higher
Energy of one photon (electron volts)
10^-9 10^-8 10^-7 10^-6 10^-5 10^-4 10^-3 10^-2 10^-1 1 10^1 10^2 10^3 10^4 10^5 10^6

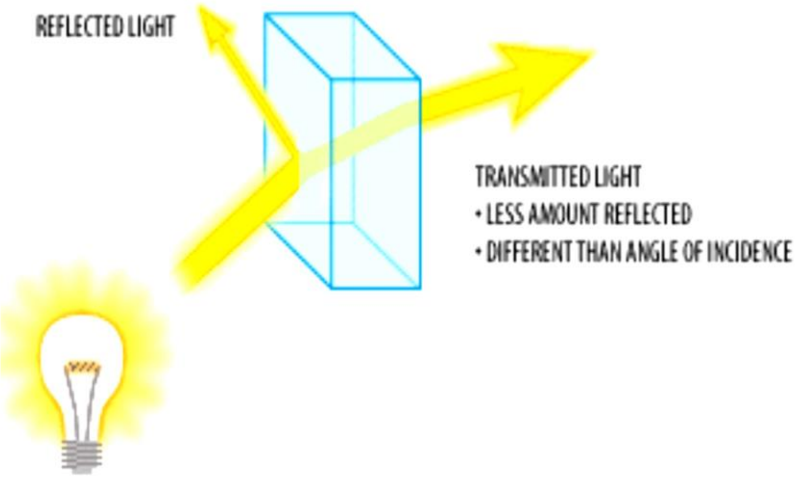
Human visual system

- Visible wavelengths: 380-780 nanometers.
- Why do we see the visible spectrum and not other frequencies of light?
  - Rhodopsins, photopsins, melanopsins the biological chemicals that transduce light in humans, only respond at these wavelengths.
- **Cones** are sensitive to color.
- **Rods** give a general, overall picture of the field of view and are not involved in color vision.
- Sensitive to low levels of illumination.



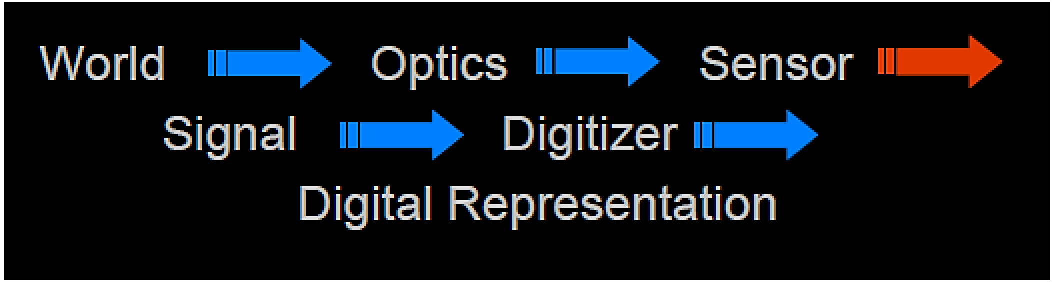
Interaction of Light and Matter

- When light strikes an object,
  - It will be wholly or partly transmitted.
  - It will be wholly or partly reflected.
  - It will be wholly or partly absorbed.
  - Physical surface properties dictate
- what happens when we see an object as blue or red or purple,
  - what we're really seeing is a partial reflection of light from that object.
  - The color we see is what's left of the spectrum after part of it is absorbed by the object.



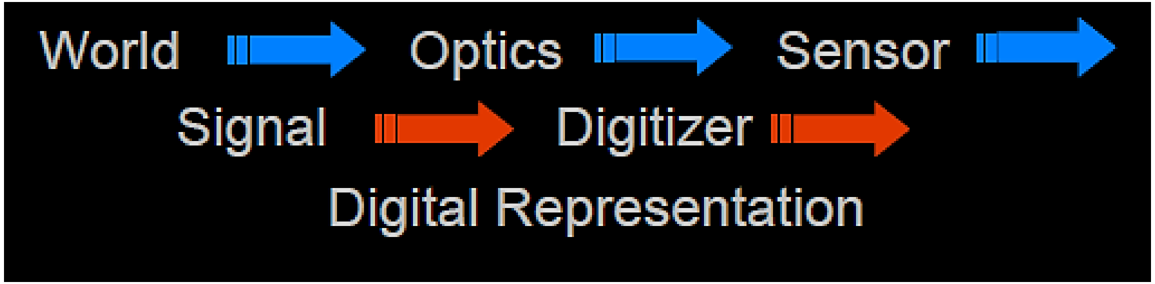
Photometry

- Photometry is the science of the measurement of light, in terms of its perceived brightness to the human eye.
- It is distinct from radiometry, which is the science of measurement of radiant energy (including light) in terms of absolute power.
- Concerned with mechanisms for converting light energy into electrical energy.



Digitization

- Digitization: conversion of the continuous (in space and value) electrical signal into a digital signal (digital image)
- It includes **Sampling** and **Quantization**.



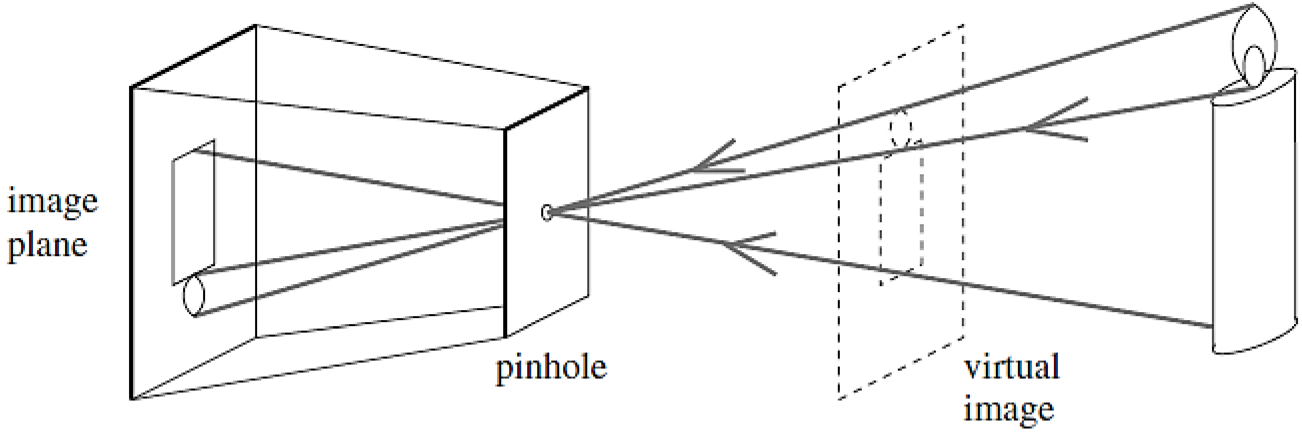
Cameras and Projections

Basics of optics

- Two models are commonly used:
  - Pin-hole camera
  - Optical system composed of lenses

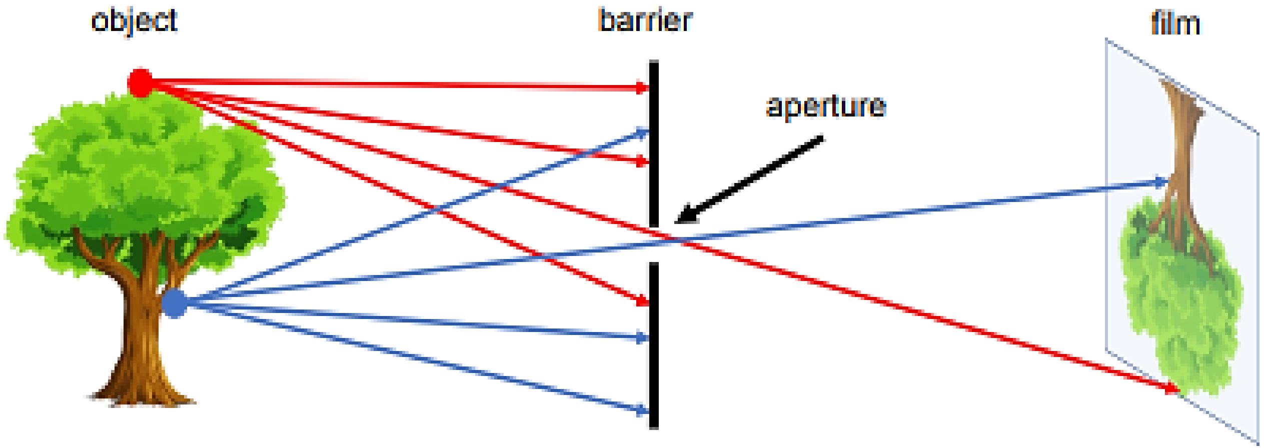
Pinhole camera model

- The pinhole camera is the simplest kind of camera. **It does not have a lens**. It just makes use of a tiny opening (a pinhole-sized opening) to focus all light rays within the smallest possible area to obtain an image, as clearly as possible. The simple image formed using a pinhole camera is always inverted

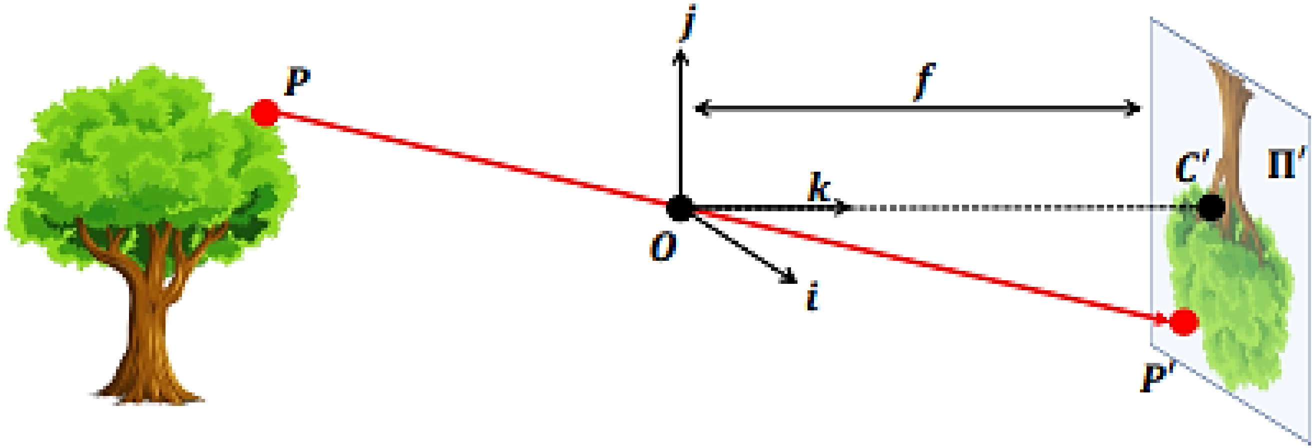


Pinhole camera model

- The simplest device to form an image of a 3D scene on a 2D surface is the "pinhole“ camera.
- Rays of light pass through a "pinhole" and form an inverted image of the object on the image plane.
- World projected to 2D Image
  - Image inverted
  - Size reduced
  - Image is dim
  - No direct depth information
- Known as **perspective projection**
- Optical Axis: the perpendicular from the image plane through the pinhole (also called the central projection ray)
- Each point in the image corresponds to a particular direction defined by a ray from that point through the pinhole.

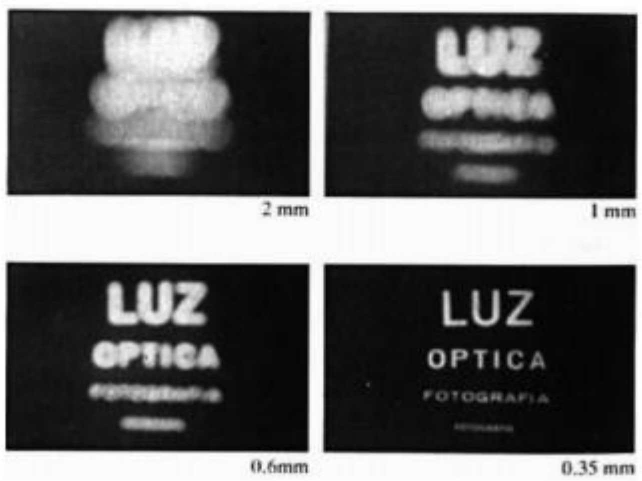


- A simple pin hole camera model



- A formal construction of the pinhole camera model

- In this construction, the film is commonly called the **image or retinal plane**. The aperture is referred to as the **pinhole** O or center of the camera. The distance between the image plane and the pinhole O is the **focal length f**. Sometimes, the retinal plane is placed between O and the 3D object at a distance f from O.

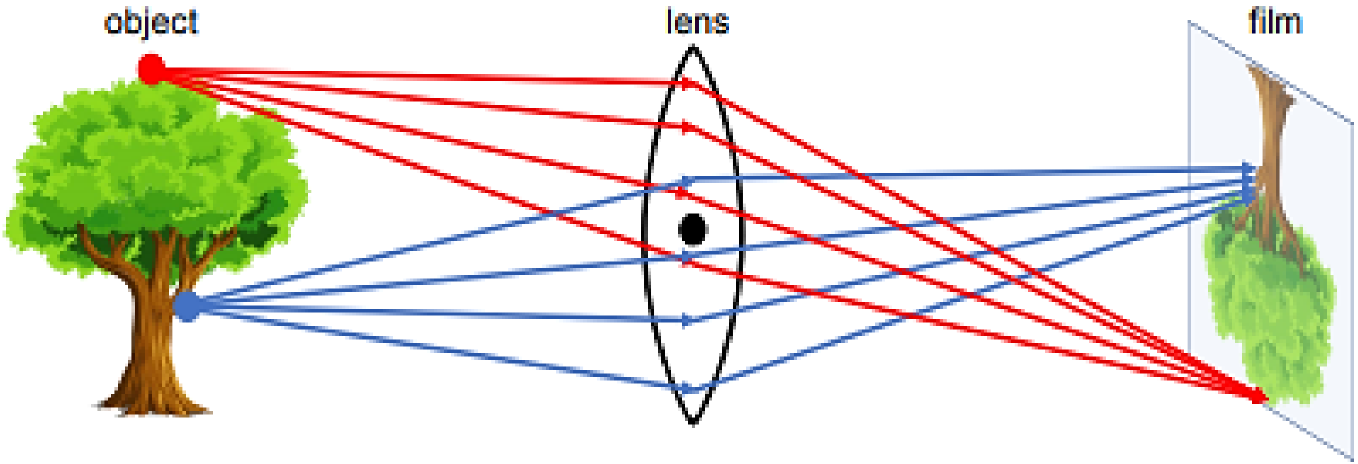


- The effects of aperture size on the image. As the aperture size decreases, the image gets sharper, but darker

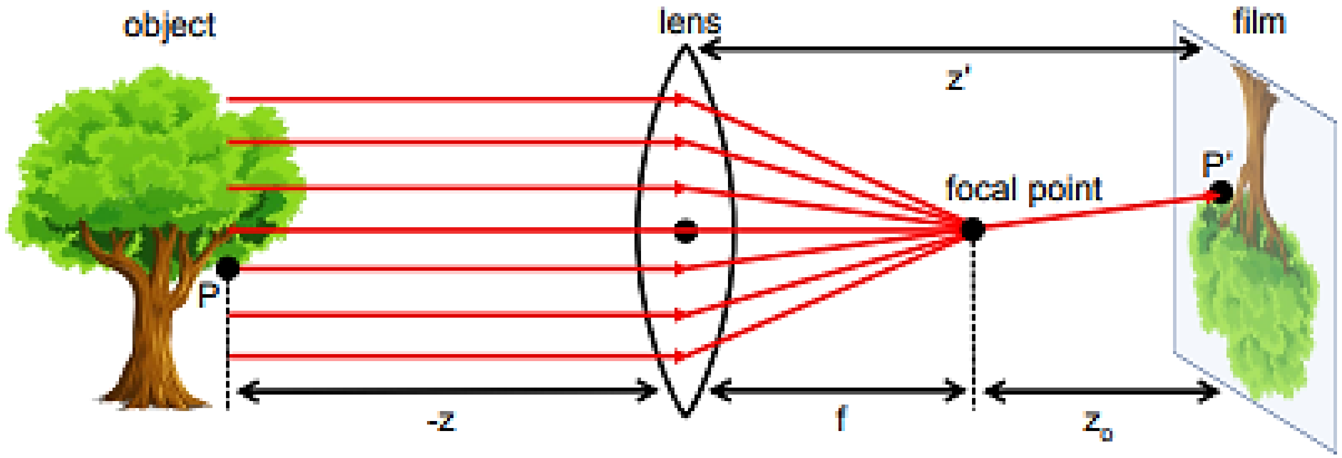
- we arrive at the fundamental problem presented by the pinhole formulation: can we develop cameras that take crisp and bright images?

Lenses

- In modern cameras, the above conflict between crispness and brightness is mitigated by using lenses, devices that can focus or disperse light.
- If we replace the pinhole with a lens that is both properly placed and sized, then it satisfies the following property: all rays of light that are emitted by some point P are refracted by the lens such that they converge to a single point P’



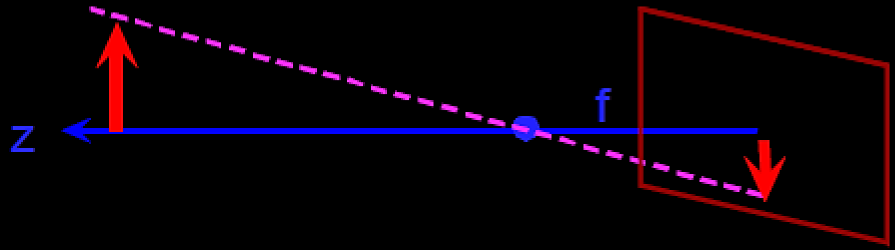
- A setup of a simple lens model. Notice how the rays of the top point on the tree converge nicely on the film. However, a point at a different distance away from the lens results in rays not converging perfectly on the film.



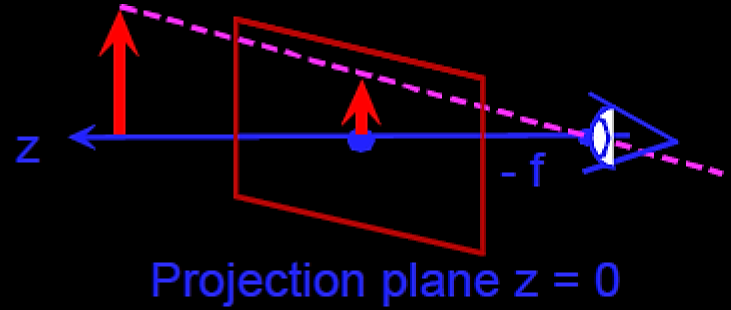
- Lenses focus light rays parallel to the optical axis into the focal point.

Pin hole camera equivalent geometry

- Consider case with object on the optical axis:

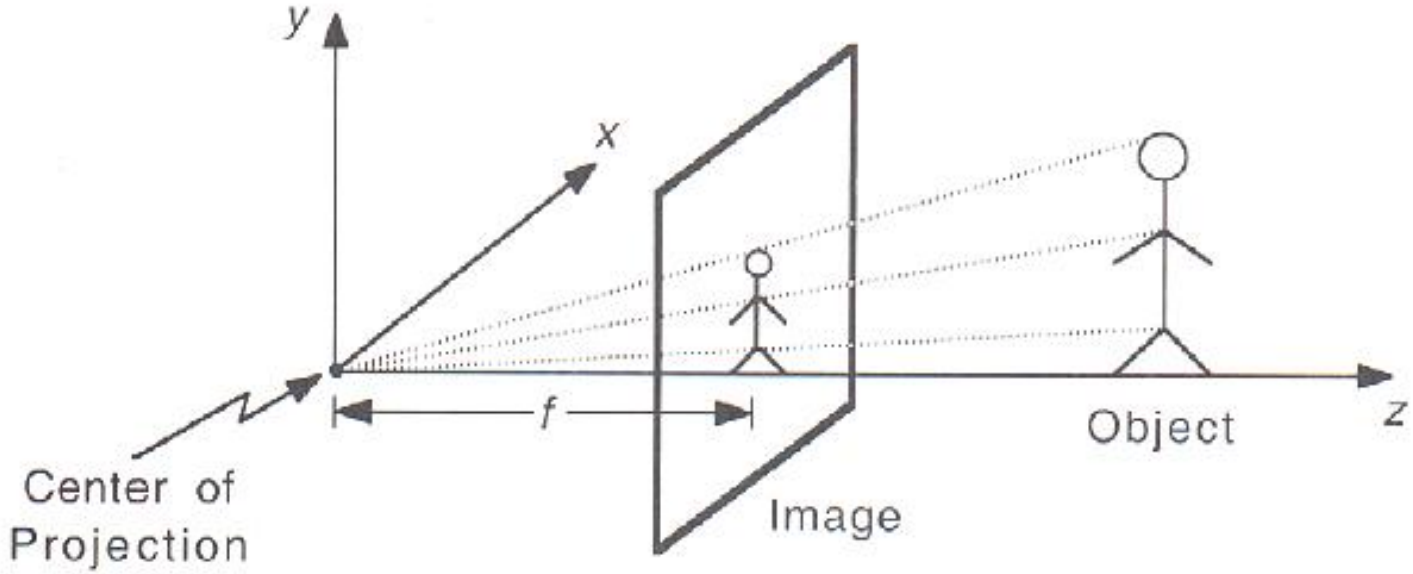


- More convenient with upright image:



- Equivalent mathematically

Perspective projection



Perspective projection

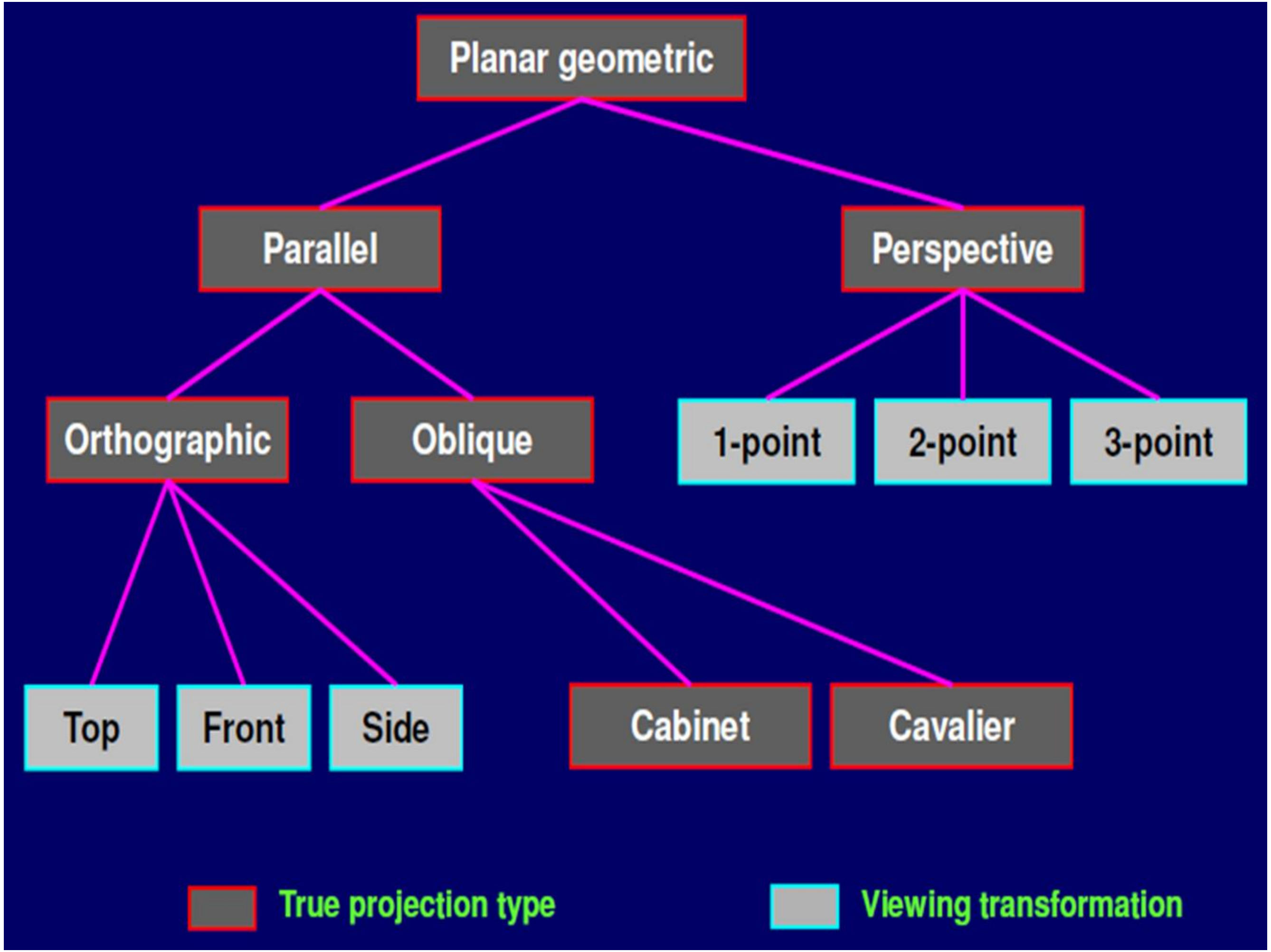
$$X_i = \frac{f}{z_0} x_0$$

$$Y_i = \frac{f}{z_0} y_0$$

$$(x_0, y_0, z_0) \rightarrow (X_i, Y_i)$$

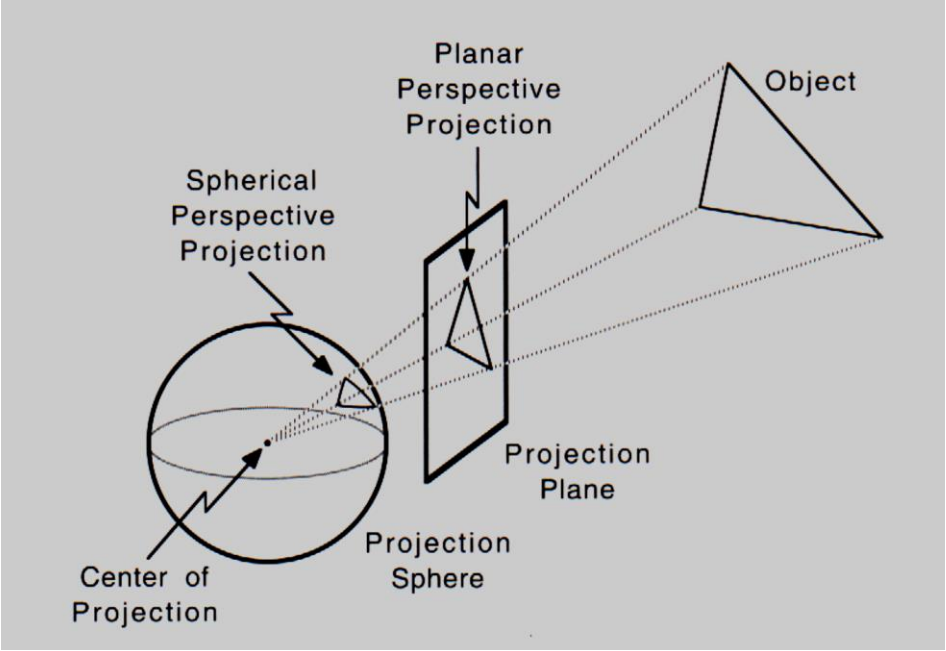
3D real world coordinates

2D image plane coordinates


Planar geometric
Parallel
Perspective
Orthographic
Oblique
1-point
2-point
3-point
Top
Front
Side
Cabinet
Cavalier
True projection type
Viewing transformation

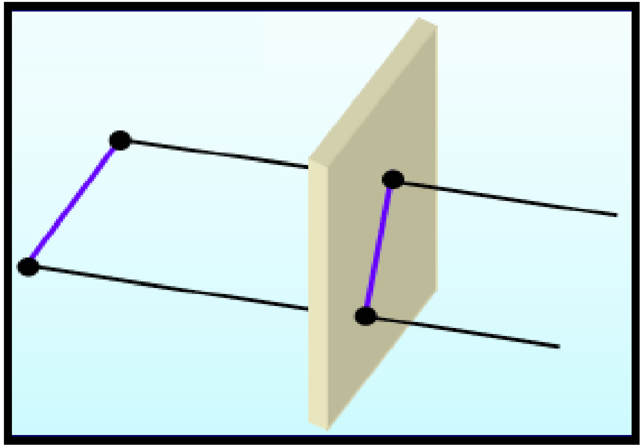
Spherical Perspective Projection

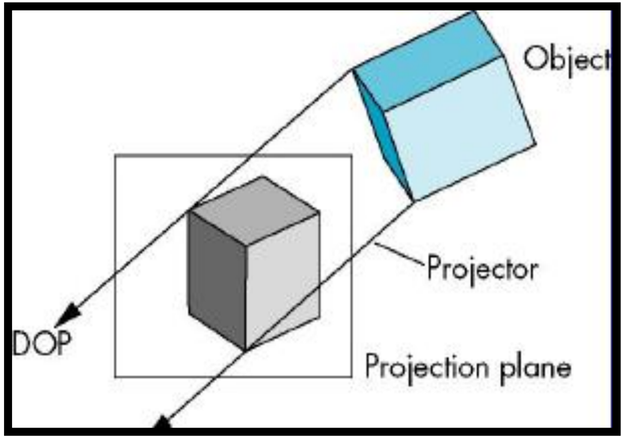
- A spherical projection **shows where lines or planes that intersect the surface of a (hemi)sphere**, provided that the lines/planes also pass through the center of the (hemi)sphere.



Parallel Projection

- Extending parallel lines from each vertex of the object until they intersect the plane of the screen
- Connect projected vertices by line segment
- To represent a 3 D object on 2D plane, simple way to discard z co ordinate.





Parallel Projection

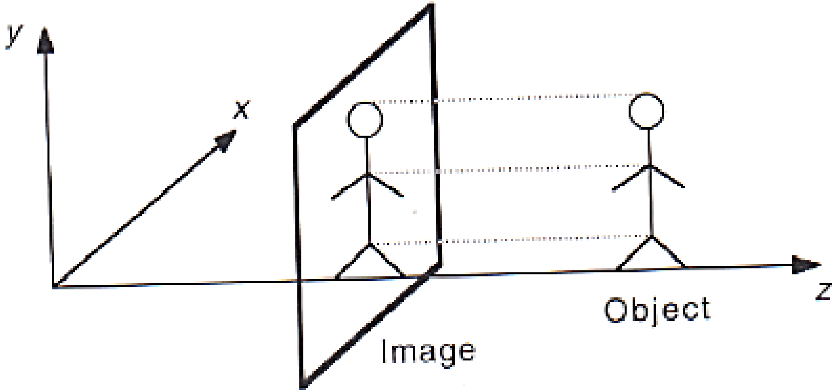
- Parallel projection is a kind of projection where the projecting lines emerge parallelly from the polygon surface and then incident parallelly on the plane.
- **Parallel projection is further divided into two categories:**
  - Orthographic Projection - It is a kind of parallel projection where the projecting lines emerge parallelly from the object surface and incident perpendicularly at the projecting plane.
  - Oblique Projection - It is a kind of parallel projection where projecting rays emerges parallelly from the surface of the polygon and incident at an angle other than 90 degrees on the plane.

Orthographic Projection

- Projection onto a plane by a set of parallel rays orthogonal to this plane.

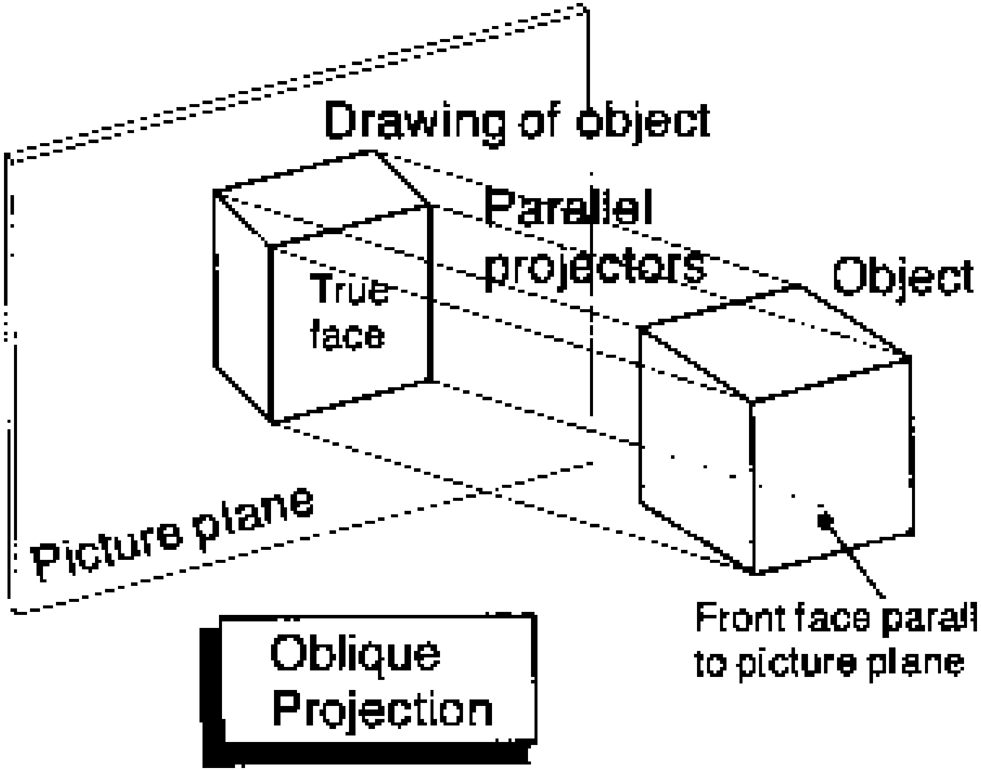
$$X_i = x_0$$

$$Y_i = y_0$$



(source: A Guided tour of computer vision/Vic Nalwa)

Oblique Projection



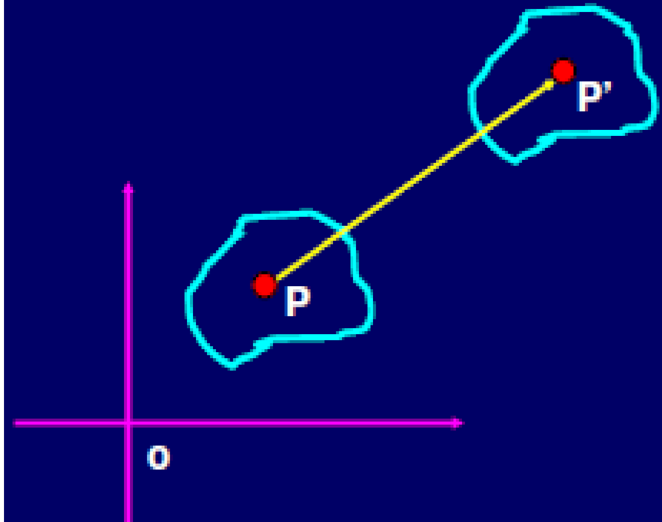
rigid and affine transformation

Geometric Transformation:

Single coordinate system.

After transformation, new coordinate of object can be obtained from original coordinates.

Object will move, viewer is fixed



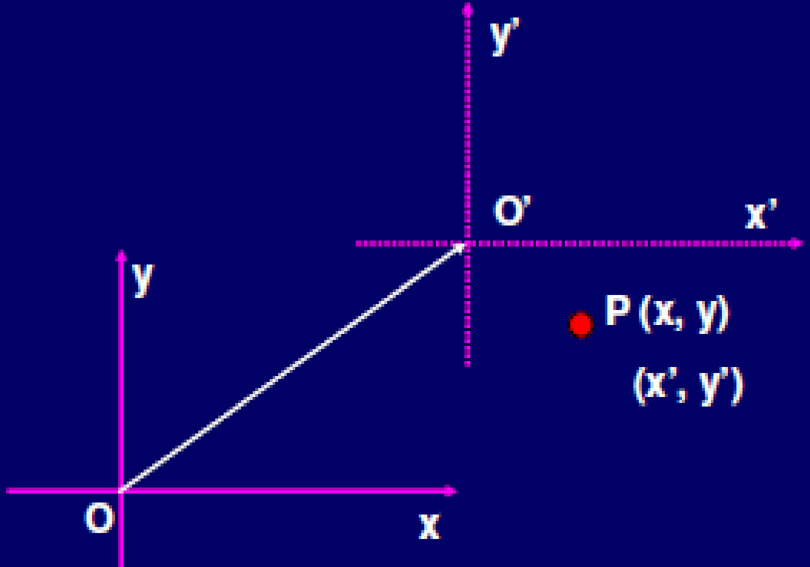
$$X' = X + Tx$$

$$Y' = Y + Ty$$

Geometric Transformation:

More than one coordinate system.

Object is stationary, Viewer will move



$$X' = X - Tx$$

$$Y' = Y - Ty$$

Matrix representation

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \bullet \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} Ax + By \\ Cx + Dy \end{bmatrix}$$

Column Major

Row Major

$$[\,x \;\; y\,] \begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} Ax + Cy & Cx + Dy \end{pmatrix}$$

We'll use the column-vector representation for a point.

2 D Geometric Transformations

- Basic 2D geometric transformations
  - 2D translation
  - 2D rotation
  - 2D scaling
- 2D Composite transformations
- A general rigid body transformation
- Other 2D geometric transformations
  - Reflection
  - Shear

Translation

A translation moves all points in an object along the same straight-line path to new positions.

The path is represented by a vector, called the translation or shift vector.

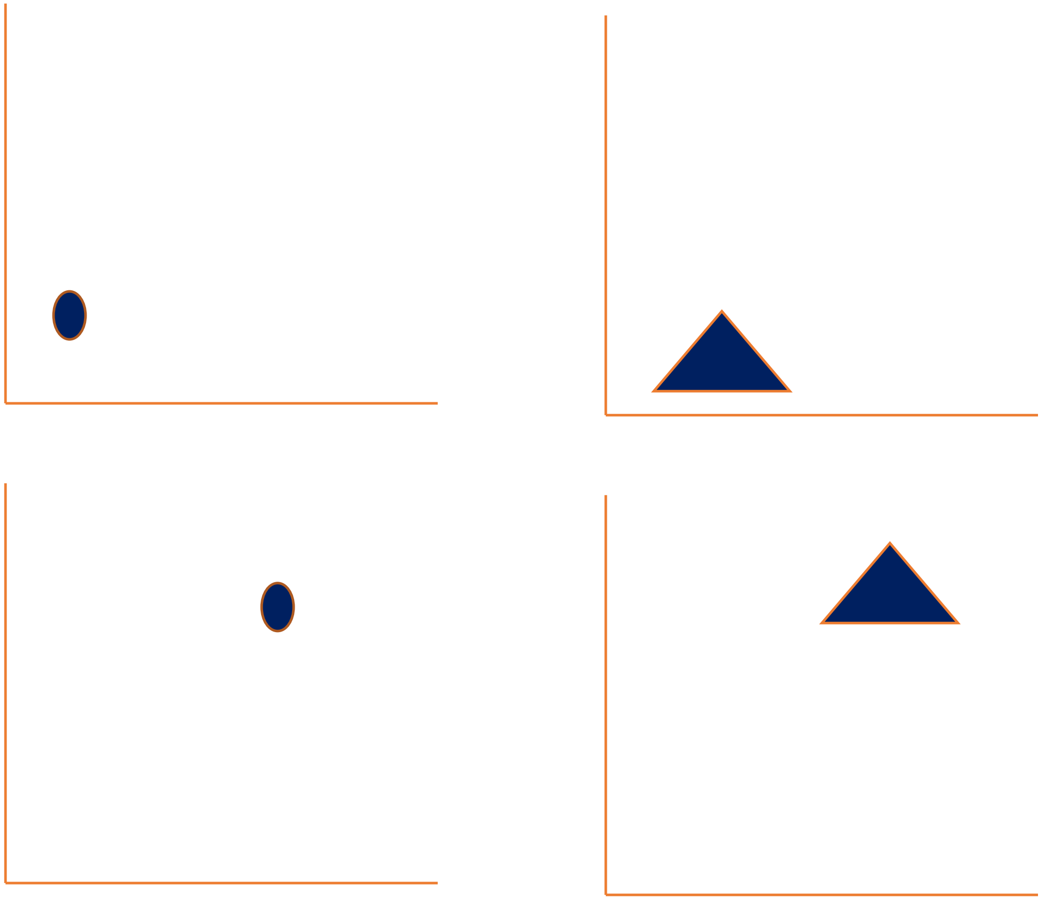
Equation:

$$x' = x + t_x$$
$$y' = y + t_y$$

$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad P = \begin{bmatrix} x \\ y \end{bmatrix}, \quad T = \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

$$P' = P + T$$

Matrix



Rotation

A rotation repositions all points in an object along a **circular path** in the plane centered at the **pivot point**.

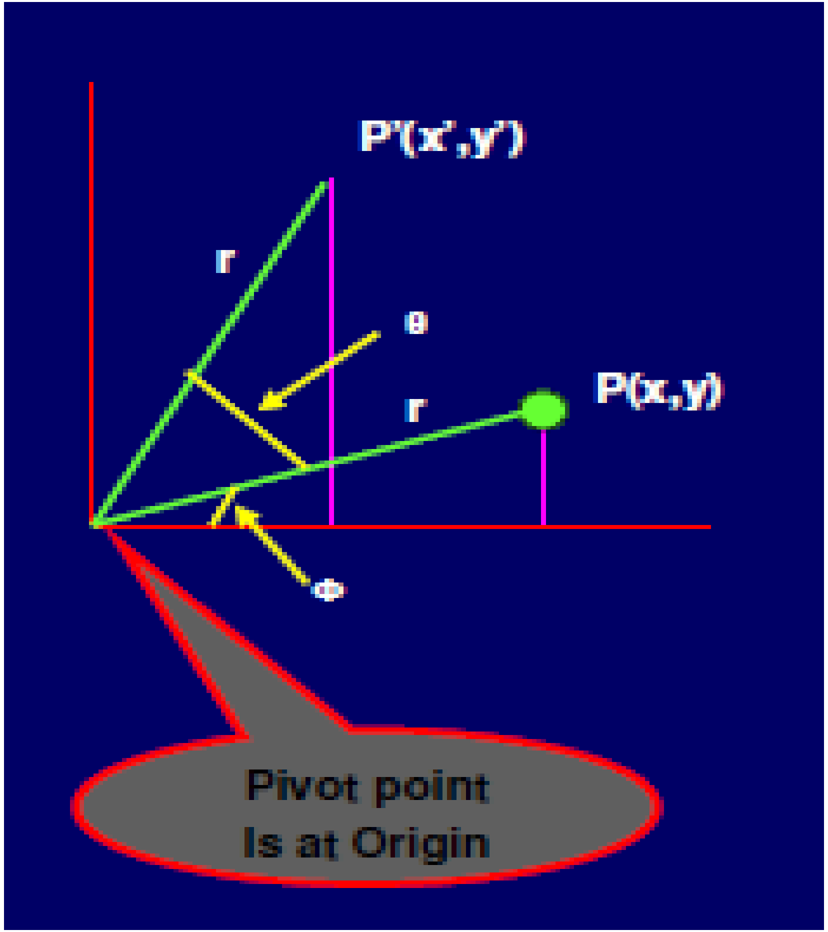
First, we'll assume the pivot is at the origin.

We can write the components:

$$p'_x = p_x \cos\theta - p_y \sin\theta$$
$$p'_y = p_x \sin\theta + p_y \cos\theta$$

or in matrix form:

$$P' = R \cdot P$$



Rotation

Equation:

$$x' = r\cos(\phi+\theta) = r\cos\phi\cos\theta - r\sin\phi\sin\theta$$
$$= x\cos\theta - y\sin\theta$$
$$y' = r\sin(\phi+\theta) = r\sin\phi\cos\theta + r\cos\phi\sin\theta$$
$$= y\cos\theta + x\sin\theta$$

Matrix:

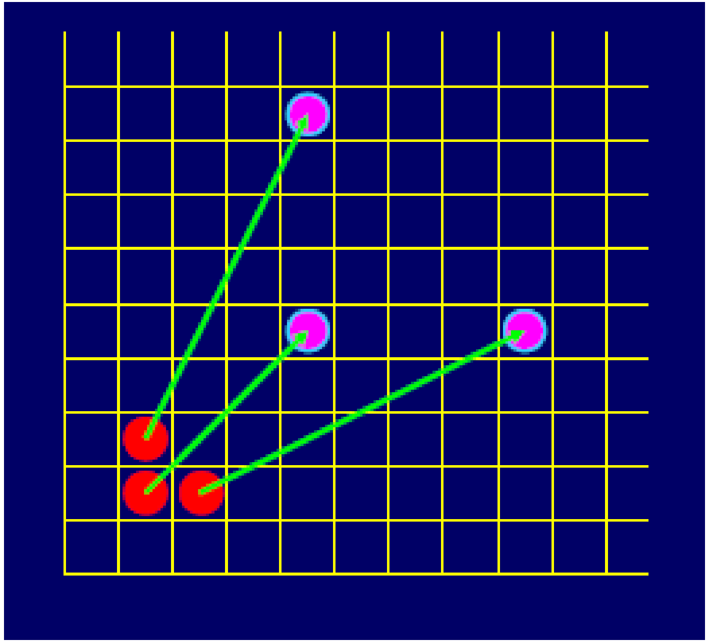
$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad P = \begin{bmatrix} x \\ y \end{bmatrix} \quad \mathrm{R} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$P' = \mathrm{R}\,P$$

A positive value for the angle defines a counter-clockwise rotation about the pivot point, and a negative value rotates objects in the clockwise direction.

Scaling

Scaling alters the size of an object. Scales are about the origin.



$$x' = s_x \cdot x$$

$$y' = s_y \cdot y$$

$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad P = \begin{bmatrix} x \\ y \end{bmatrix}, \quad S = \begin{bmatrix} S_x & 0 \\ 0 & Sy \end{bmatrix}$$

$$P' = SP$$

Combining transformations

We have a general transformation of a point:

$P' = M \cdot P + A$

When we scale or rotate, we set M, and A is the additive identity (0).

When we translate, we set A, and M is the multiplicative identity (1).

To combine multiple transformations, we must explicitly compute each transformed point.

Inverse transformation

Inverse Translation Matrix :

$$T^{-1} = \begin{bmatrix} 1 & 0 & -tx \\ 0 & 1 & -ty \\ 0 & 0 & 1 \end{bmatrix}$$

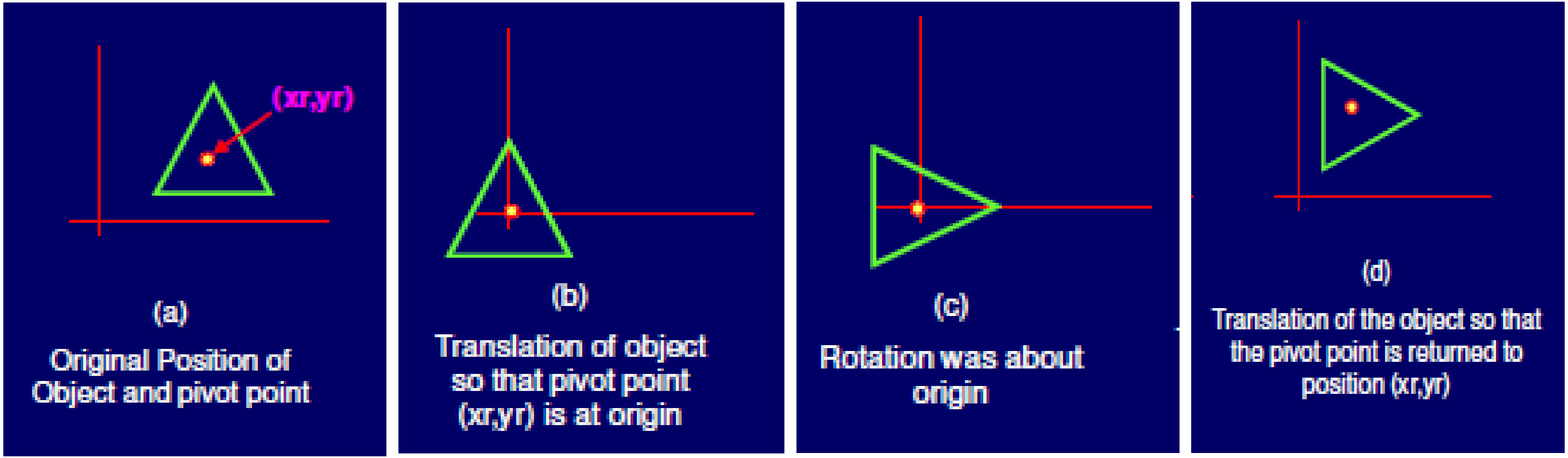
Inverse Rotation Matrix :

$$R^{-1} = \begin{bmatrix} \cos(-\theta) & -\sin(-\theta) & 0 \\ \sin(-\theta) & \cos(-\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Inverse Scaling Matrix :

$$S^{-1} = \begin{bmatrix} \frac{1}{s_x} & 0 & 0 \\ 0 & \frac{1}{s_y} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

General Pivot Point Rotation



General Pivot Point Rotation

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -x_r \\ -y_r \end{bmatrix}$$

(1) Translation

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

(2) Rotation

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x''' \\ y''' \end{bmatrix} + \begin{bmatrix} x_r \\ y_r \end{bmatrix}$$

(3) Translation

$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix}, \quad P = \begin{bmatrix} x \\ y \end{bmatrix},$$

$$T_1 = \begin{bmatrix} -x_r \\ -y_r \end{bmatrix} \quad R = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}, \quad T_2 = \begin{bmatrix} x_r \\ y_r \end{bmatrix}$$

Composite Transformation

$$P' = R(P + T_1) + T_2 = RP + (RT_1 + T_2)$$

Rotation about General Point

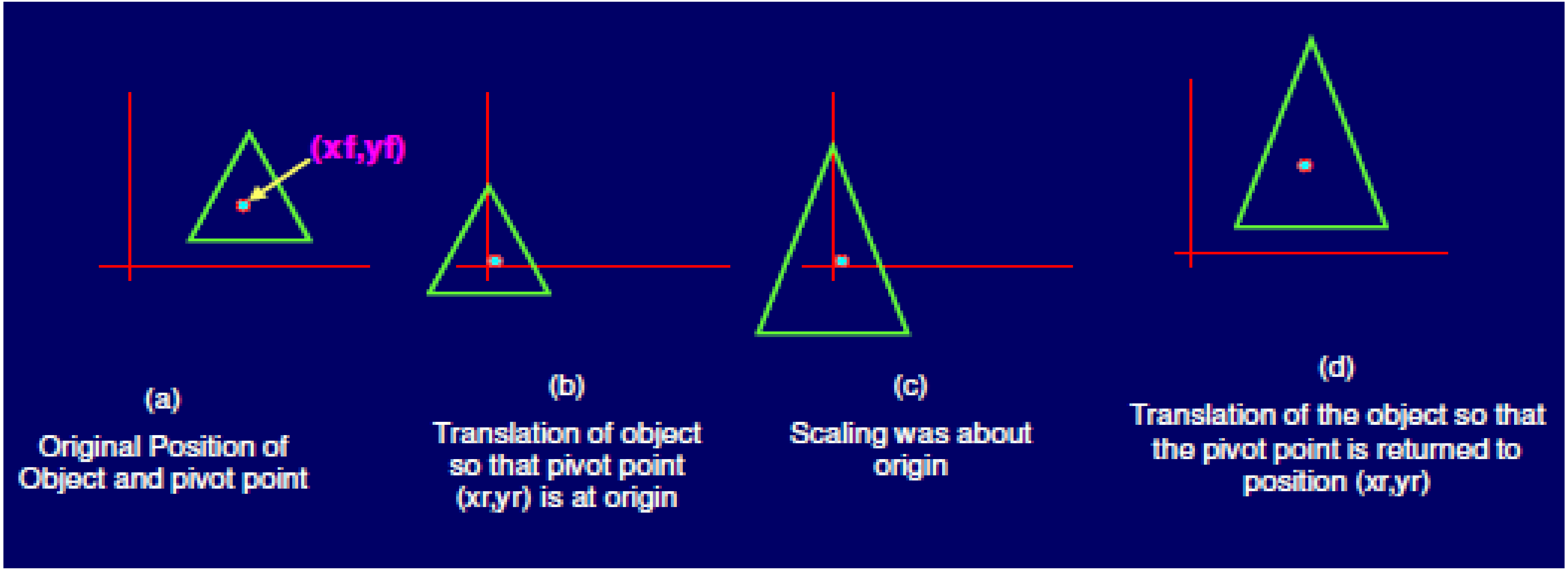
$$\begin{bmatrix} 1 & 0 & x_r \\ 0 & 1 & y_r \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -x_r \\ 0 & 1 & -y_r \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos\theta & -\sin\theta & xr(1-\cos\theta) + yr\sin\theta \\ \sin\theta & \cos\theta & yr(1-\cos\theta) - xr\sin\theta \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix

Can also be expressed as $T(xr,yr).\,R(\theta).T(-xr,-yr) = R(xr,yr,\theta)$

General Fixed Point Scaling



Fixed Point Scaling

$$\begin{bmatrix} x'' \\ y'' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} -x_f \\ -y_f \end{bmatrix}$$

(1) Translation

$$\begin{bmatrix} x''' \\ y''' \end{bmatrix} = \begin{bmatrix} Sx & 0 \\ 0 & Sy \end{bmatrix} \begin{bmatrix} x'' \\ y'' \end{bmatrix}$$

(2) Scaling

$$P' = \begin{bmatrix} x' \\ y' \end{bmatrix} \quad P = \begin{bmatrix} x \\ y \end{bmatrix}$$

$$T_1 = \begin{bmatrix} -x_f \\ -y_f \end{bmatrix} \quad R = \begin{bmatrix} Sx & 0 \\ 0 & Sy \end{bmatrix} \quad T_2 = \begin{bmatrix} x_f \\ y_f \end{bmatrix}$$

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x''' \\ y''' \end{bmatrix} + \begin{bmatrix} x_f \\ y_f \end{bmatrix}$$

(3) Translation

Composite Transformation

$$P' = S(P + T_1) + T_2 = SP + (ST_1 + T_2)$$

Scaling with respect to fixed point

$$\begin{bmatrix} 1 & 0 & xf \\ 0 & 1 & yf \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} Sx & 0 & 0 \\ 0 & Sy & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & -xf \\ 0 & 1 & -yf \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} Sx & 0 & xf(1- Sx) \\ 0 & Sy & yf(1- Sf) \\ 0 & 0 & 1 \end{bmatrix}$$

Transformation matrix

Can also be expressed as T(xf,yf). S(sx,sy). T(-xf,-yf) = S(xf, yf, sx, sy)

Composite Translation

P' = t1.P

P'' = t2.P' So, P'' = t2.t1.P

$$\begin{pmatrix} 1 & 0 & t_{x2} \\ 0 & 1 & t_{y2} \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & t_{x1} \\ 0 & 1 & t_{y1} \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & t_{x1} + t_{x2} \\ 0 & 1 & t_{y1} + t_{y2} \\ 0 & 0 & 1 \end{pmatrix}$$

$$\begin{aligned} T &= \mathrm{T}(t_{x1}, t_{y1}) \cdot \mathrm{T}(t_{x2}, t_{y2}) \\ &= \mathrm{T}(t_{x1} + t_{x2}, t_{y1} + t_{y2}) \end{aligned}$$

Two Successive Translations are *Additive*

Composite Rotation

$$T = \mathbf{R}(\theta_2) \cdot \mathbf{R}(\theta_1)$$
$$= \mathbf{R}(\theta_2 + \theta_1)$$

$$\begin{pmatrix} \cos\theta_2 & -\sin\theta_2 & 0 \\ \sin\theta_2 & \cos\theta_2 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \cos\theta_1 & -\sin\theta_1 & 0 \\ \sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} \cos(\theta_2+\theta_1) & -\sin(\theta_2+\theta_1) & 0 \\ \sin(\theta_2+\theta_1) & \cos(\theta_2+\theta_1) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Two Successive Rotations are *Additive*

Composite Scaling

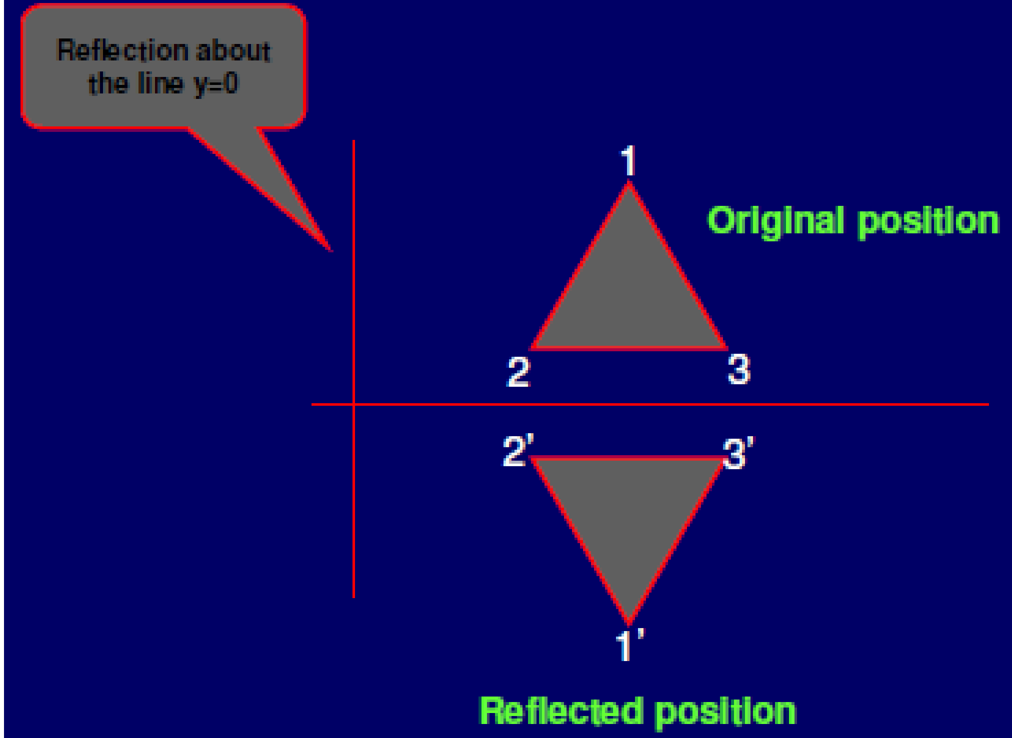
$$T = \mathbf{S}(s_{x1}, s_{y1}) \cdot \mathbf{S}(s_{x2}, s_{y2})$$
$$= \mathbf{S}(s_{x1} s_{x2}, s_{y1} s_{y2})$$

$$\begin{pmatrix} s_{x2} & 0 & 0 \\ 0 & s_{y2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} s_{x1} & 0 & 0 \\ 0 & s_{y1} & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} s_{x1} \cdot s_{x2} & 0 & 0 \\ 0 & s_{y1} \cdot s_{y2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Two Successive Scalings are *Multiplicative*

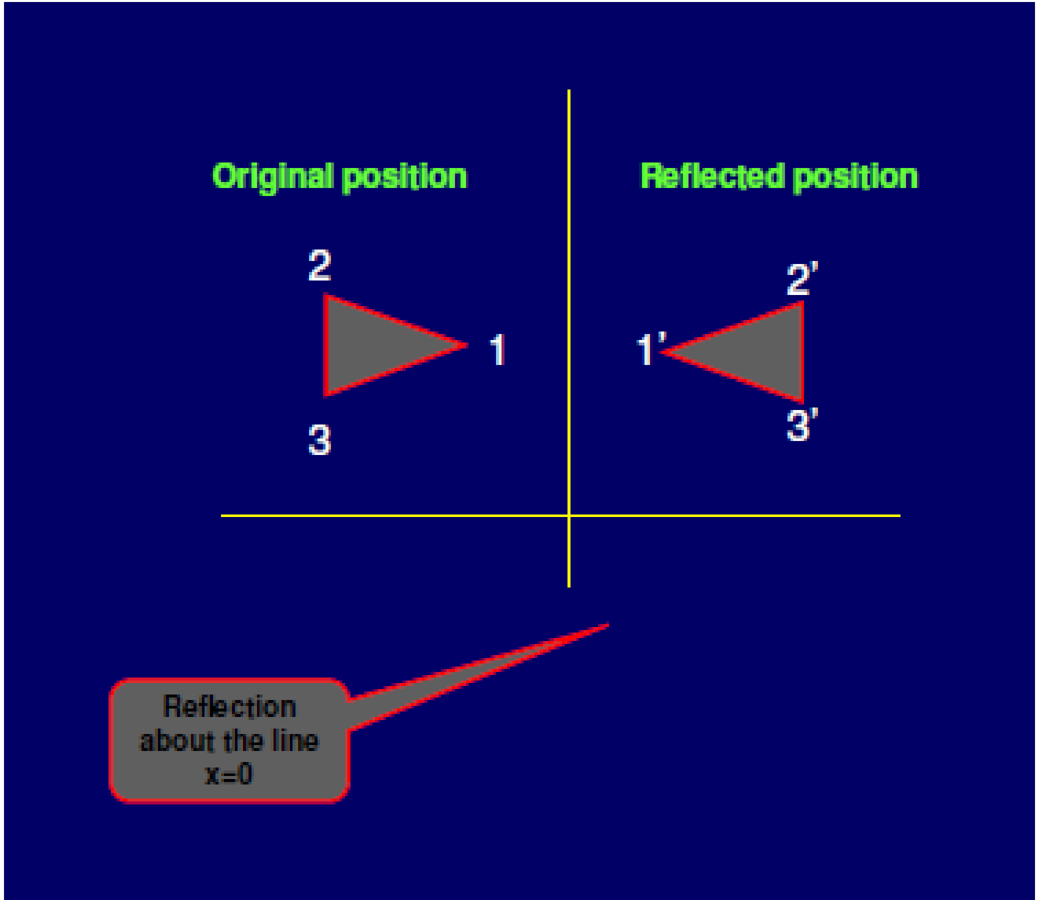
Reflection

Reflection : A transformation that produces a mirror image of an object. It is obtained by rotating the object 180° about the reflection axis



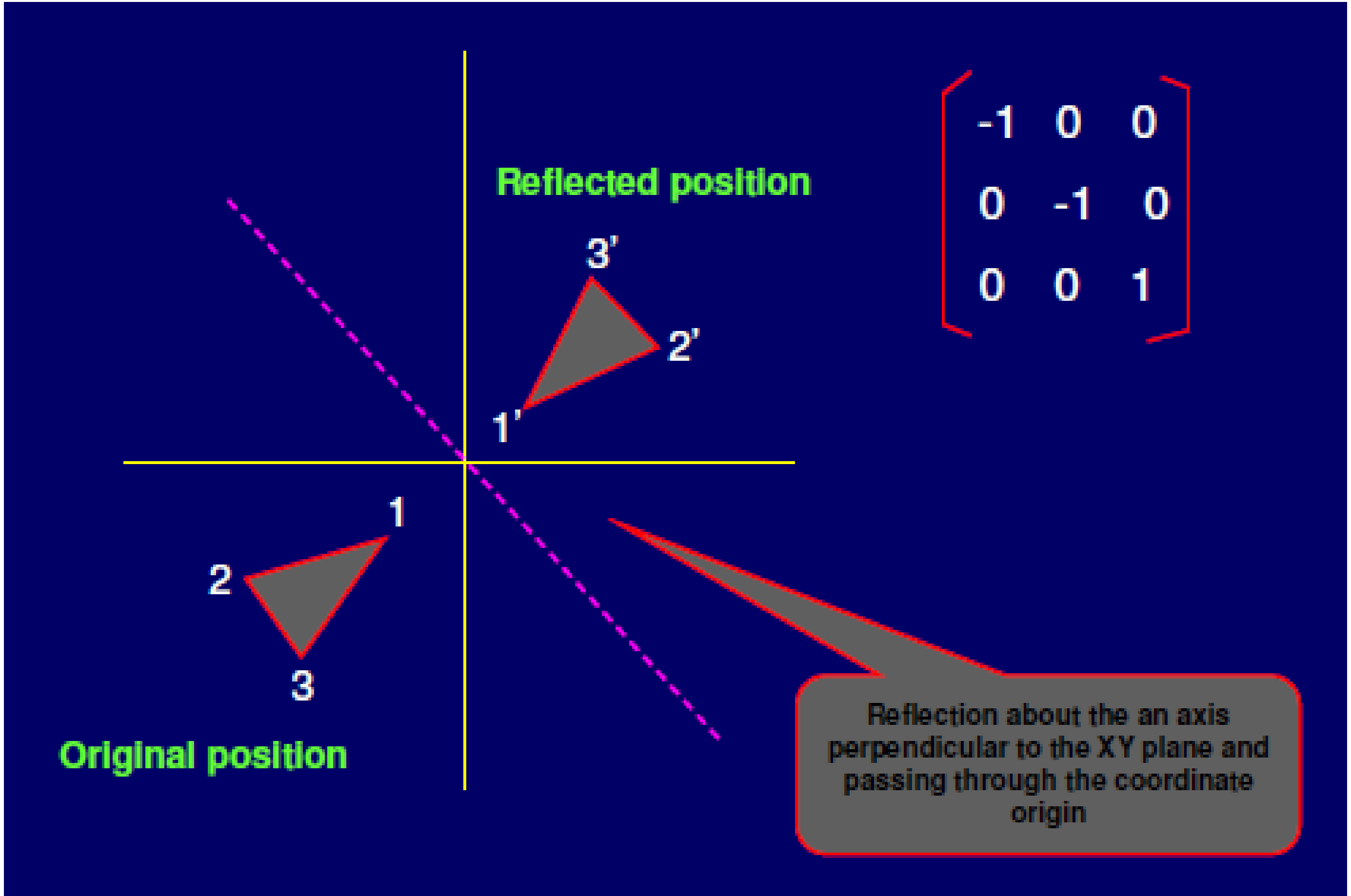
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Reflection



$$\begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Reflection



Reflected position
-1 0 0
0 -1 0
0 0 1
3'
2'
1'
1
2
3
Original position
Reflection about the an axis perpendicular to the XY plane and passing through the coordinate origin

Shear

Shearing : A transformation that distorts the shape of an object such that the transformed shape appears as if the object were composed of internal layers that had been caused to slide over each other

Two common shearing transformations are those that shift coordinate x values and those that shift y values

$$\begin{bmatrix} 1 & sh_x & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$x' = x + sh_x * y$$

$$y' = y$$

Shear in X - Direction

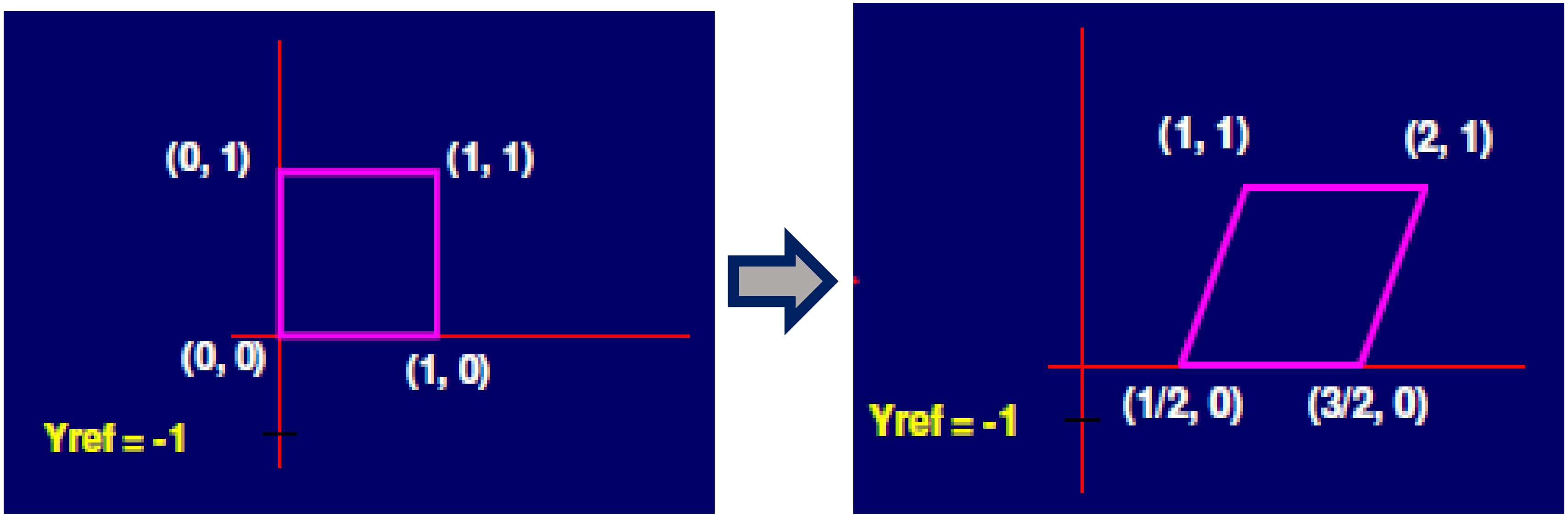
$$\begin{bmatrix} 1 & 0 & 0 \\ sh_y & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$x' = x$$

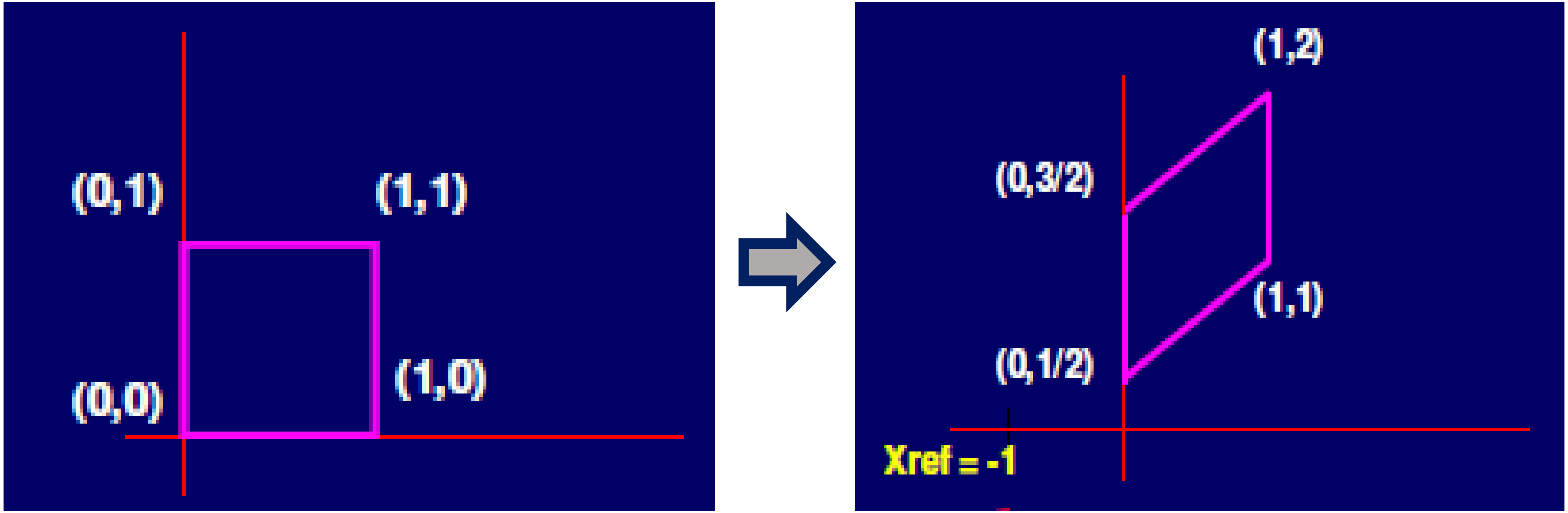
$$y' = y + sh_y * x$$

Shear in Y - Direction

X direction Shear



Y direction Shear



Rigid body and Affine Transformations

- **Rigid Body Transformation :** Those transformations which does not change the shape of object after transformation.

  Preserves length and angle

  - Translation
  - Rotation
  - Reflection

- **Affine Transformation :** Those operation which maps parallel lines into parallel lines and infinite points to infinite points.

  - Does not preserve length and angle
  - All five (Translation, Rotation, Reflection, Scaling & Shearing) transformations.