Computer Vision – Transformations, Imaging Geometry and Stereo Vision

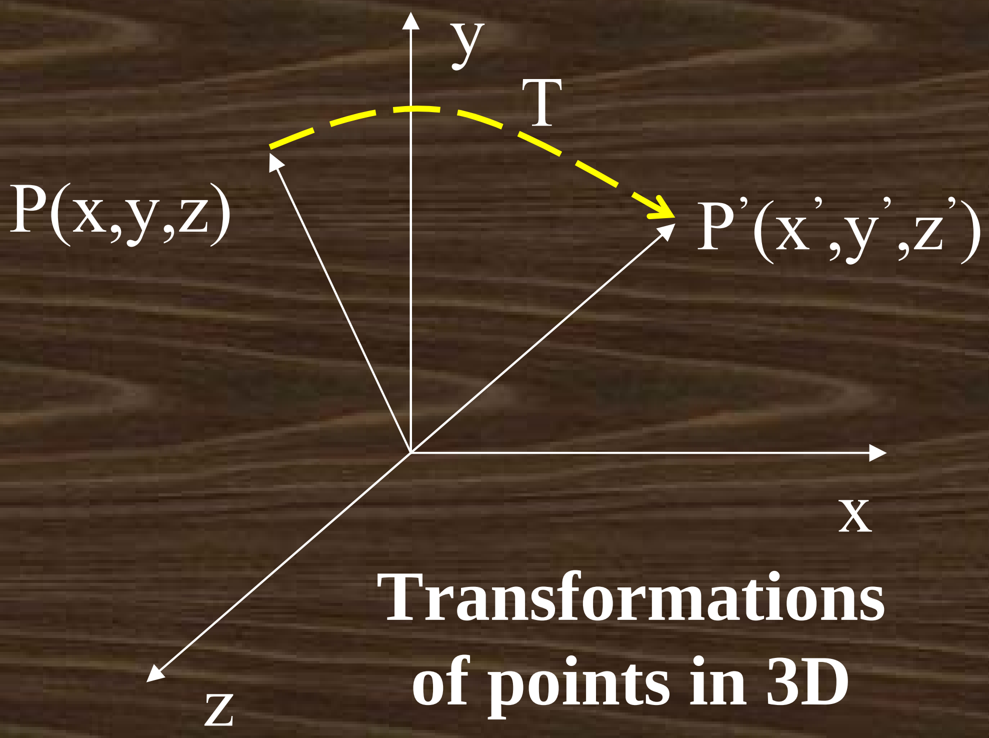
Dr. S. Das

IIT Madras, Chennai-36

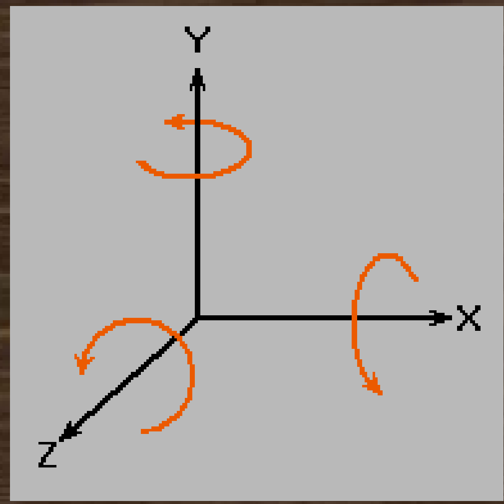
BASICS

Representation of Points in the 3D world: a vector of length 3

$$\overline{X} = [x\ y\ z]^T$$



Transformations of points in 3D



Right handed coordinate system

4 basic transformations

- **Translation**
- **Rotation**
- **Scaling**
- **Shear**

Affine transformations

Basics 3D Transformation equations

- Translation : P' = P + ΔP

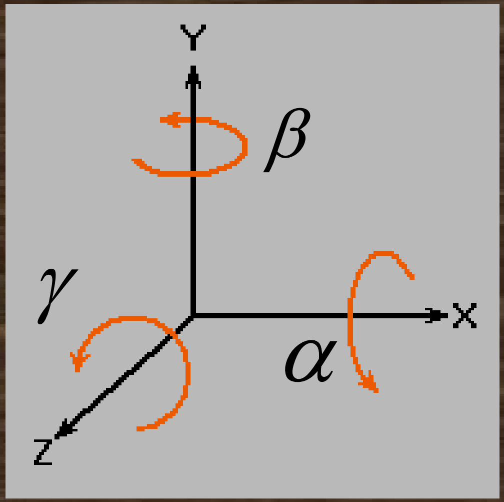
$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \\ \Delta z \end{bmatrix}$$

- Scaling: P'= SP

$$S = \begin{bmatrix} S_x & 0 & 0 \\ 0 & S_y & 0 \\ 0 & 0 & S_z \end{bmatrix}$$

- Rotation : about an axis,

  P' = RP

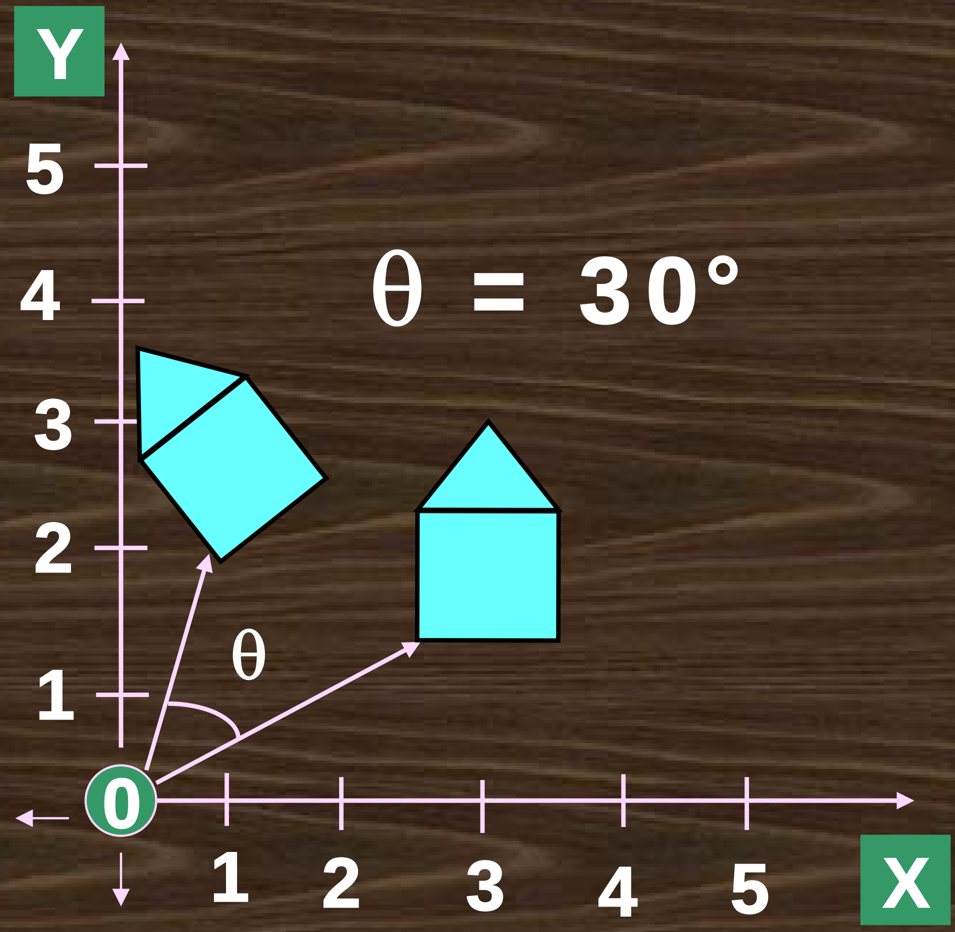


ROTATION - 2D

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

In matrix form, this is :

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



Positive Rotations: counter clockwise about the origin

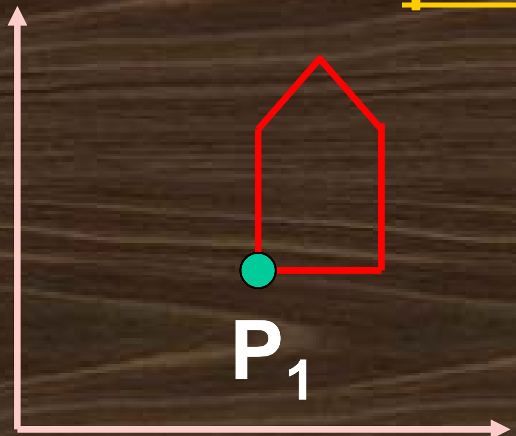
For rotations, $|R| = 1$ and $[R]^T = [R]^{-1}$. Rotation matrices are orthogonal.

Rotation about an arbitrary point P in space

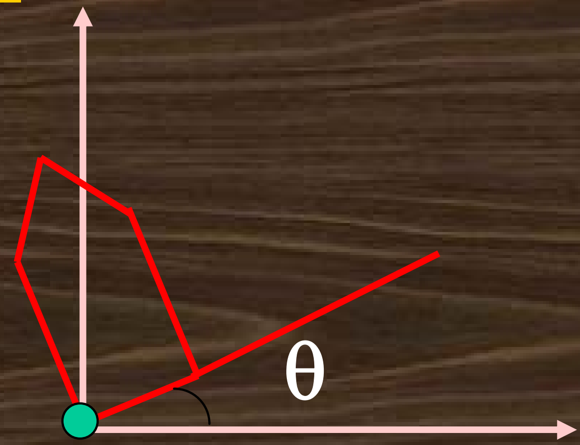
As we mentioned before, rotations are applied about the origin. So to rotate about any arbitrary point P in space, **translate** so that P coincides with the origin, then **rotate**, then **translate back**. Steps are:

- Translate by $(-P_x, -P_y)$
- Rotate
- Translate by (P_x, P_y)

Rotation about an arbitrary point P in space

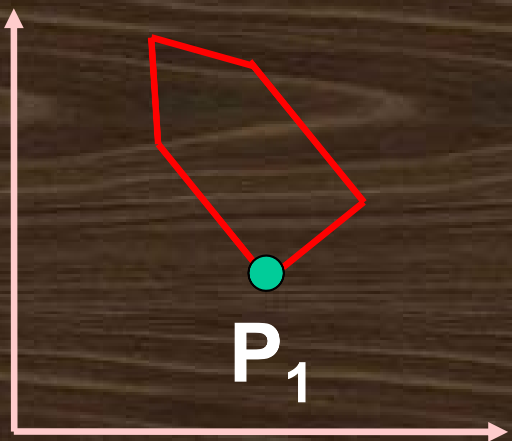


House at P_1



Rotation by θ

Translation of P_1 to Origin



Translation back to P_1

2D Transformation equations (revisited)

- Translation : P' = P + ΔP

$$\begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} \Delta x \\ \Delta y \end{bmatrix} \Longrightarrow \begin{bmatrix} x' \\ y' \end{bmatrix} = \begin{bmatrix} 1 & 0 & \Delta x \\ 0 & 1 & \Delta y \end{bmatrix}^T \begin{bmatrix} x \\ y \end{bmatrix} ??$$

- Rotation : about an axis,

  P' = RP

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

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

Rotation about an arbitrary point P in space

$$R_{gen} = T_1(-P_x, -P_y) * R_2(\theta) * T_3(P_x, P_y)$$

$$= \begin{bmatrix} 1 & 0 & -P_x \\ 0 & 1 & -P_y \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix} * \begin{bmatrix} 1 & 0 & P_x \\ 0 & 1 & P_y \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} \cos(\theta) & -\sin(\theta) & P_x * (\cos(\theta) - 1) - P_y * (\sin(\theta)) \\ \sin(\theta) & \cos(\theta) & P_y * (\cos(\theta) - 1) + P_x * \sin(\theta) \\ 0 & 0 & 1 \end{bmatrix}$$

Using Homogeneous system

Homogeneous representation of a point in 3D space:

$$P = \left| x \; y \; z \; w \right|^T$$

$$(w = 1, \text{for a 3D point})$$

Transformations will thus be represented by 4x4 matrices:

P' = A.P

Homogenous Coordinate systems

- In order to Apply a sequence of transformations to produce composite transformations we introduce the fourth coordinate
- Homogeneous representation of 3D point:

  $|x\ y\ z\ h|^T$ (h=1 for a 3D point, dummy coordinate)
- Transformations will be represented by 4x4 matrices.

$$T = \begin{bmatrix} 1 & 0 & 0 & \Delta x \\ 0 & 1 & 0 & \Delta y \\ 0 & 0 & 1 & \Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Homogenous Translation matrix

$$S = \begin{bmatrix} S_x & 0 & 0 & 0 \\ 0 & S_y & 0 & 0 \\ 0 & 0 & S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Homogenous Scaling matrix

$$R_\alpha = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation about x axis by angle α

$$R_\beta = \begin{bmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation about y axis by angle β

Change of sign?

$$R_\gamma = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 & 0 \\ \sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Rotation about z axis by angle γ

How can one do a Rotation about an arbitrary Axis in Space?

3D Transformation equations (3)

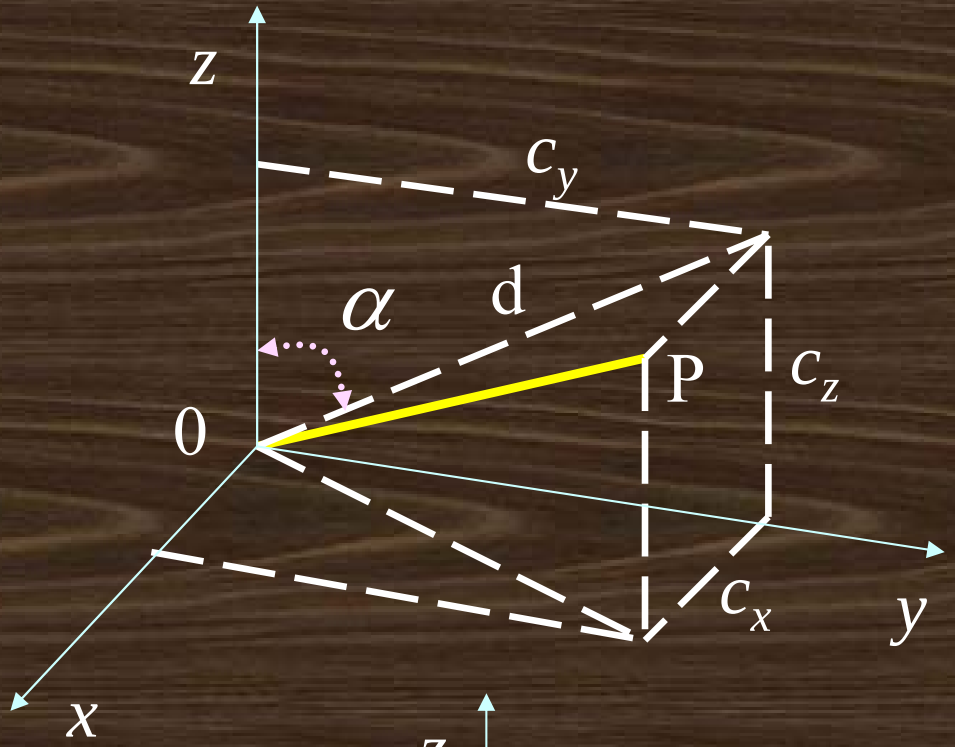
Rotation About an Arbitrary Axis in Space

Assume we want to perform a rotation about an axis in space, passing through the point *(x_0, y_0, z_0)* with direction cosines *(c_x, c_y, c_z)*, by θ degrees.

1) **First of all, translate by: - *$(x_0, y_0, z_0) = |T|$*.**
2) **Next, we rotate the axis into one of the principle axes. Let's pick, *Z ($|R_x|$, $|R_y|$)*.**
3) **We rotate next by θ degrees in *Z ($|R_z(\theta)|$)*.**
4) **Then we undo the rotations to align the axis.**
5) **We undo the translation: translate by *(x_0, y_0, z_0)***

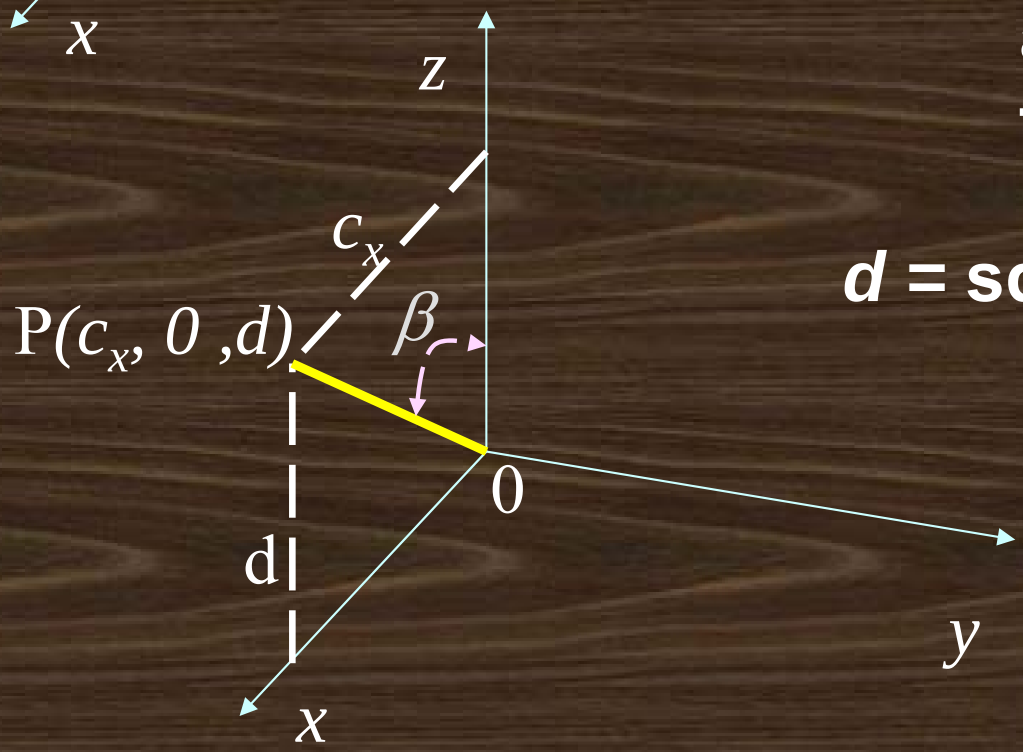
The tricky part is (2) above.

This is going to take 2 rotations,
i) about x (to place the axis in the x-z plane)
and
ii) about y (to place the result coincident with the z axis).



Rotation about *x* by α:
How do we determine α?

Project the unit vector, along OP, into the y-z plane. The y and z components are c_y and c_z, the directions cosines of the unit vector along the arbitrary axis. It can be seen from the diagram above, that :



$$d = \text{sqrt}(c_y^2 + c_z^2), \qquad \cos(\alpha) = c_z/d$$

$$\sin(\alpha) = c_y/d$$

Rotation by β about y:
How do we determine β?
Similar to above:

Determine the angle β to rotate the result into the Z axis:

The *x* component is c_x and the *z* component is d.

$$\cos(\beta) = d = d\,/(\text{length of the unit vector})$$

$$\sin(\beta) = c_x = c_x\,/(\text{length of the unit vector}).$$

Final Transformation:

$$M = |T|^{-1}\,|R_x|^{-1}\,|R_y|^{-1}\,|R_z|\,|R_y|\,|R_x|\,|T|$$

If you are given 2 points instead, you can calculate the direction cosines as follows:

$$V = |\,(x_1 - x_0)\ \ (y_1 - y_0)\ \ (z_1 - z_0)\,|^T$$

$$c_x = (x_1 - x_0)/\,|V|$$

$$c_y = (y_1 - y_0)/\,|V|$$

$$c_z = (z_1 - z_0)/\,|V|,$$

where |V| is the length of the vector V.

Inverse transformations

$$T^{-1} = \begin{bmatrix} 1 & 0 & 0 & -\Delta x \\ 0 & 1 & 0 & -\Delta y \\ 0 & 0 & 1 & -\Delta z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Inverse Translation

$$S^{-1} = \begin{bmatrix} 1/S_x & 0 & 0 & 0 \\ 0 & 1/S_y & 0 & 0 \\ 0 & 0 & 1/S_z & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Inverse scaling

Inverse Rotation

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & \sin\alpha & 0 \\ 0 & -\sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} \cos\beta & 0 & -\sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ \sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} \cos\gamma & \sin\gamma & 0 & 0 \\ -\sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

R_α^{-1} R_β^{-1} R_γ^{-1}

Concatenation of transformations

- The 4 X 4 representation is used to perform a sequence of transformations.
- Thus application of several transformations in a particular sequence can be presented by a single transformation matrix

$$v^* = R_\theta(S(Tv)) = Av; \;\; \mathrm{A} = \mathrm{R}_\theta.S.T$$

- The order of application is important… the multiplication may not be commutable.

Commutivity of Transformations

If we scale, then translate to the origin, and then translate back, is that equivalent to translate to origin, scale, translate back?

When is the order of matrix multiplication unimportant?

When does $T_1 * T_2 = T_2 * T_1$?

Cases where $T_1 * T_2 = T_2 * T_1$:

T_1	T_2
translation	translation
scale	scale
rotation	rotation
Scale (uniform)	rotation

COMPOSITE TRANSFORMATIONS

If we want to apply a series of transformations T_1, T_2, T_3 to a set of points, We can do it in two ways:

1) We can calculate $p'=T_1*p$, $p''= T_2*p'$, $p'''=T_3*p''$
2) Calculate $T= T_1*T_2*T_3$, then $p'''= T*p$.

Method 2, saves large number of additions and multiplications (computational time) – needs approximately 1/3 of as many operations. Therefore, we concatenate or compose the matrices into one final transformation matrix, and then apply that to the points.

Spaces

Object Space

definition of objects. Also called Modeling space.

World Space

where the scene and viewing specification is made

Eye space (Normalized Viewing Space)

where eye point (COP) is at the origin looking down the Z axis.

3D Image Space

A 3D Perspected space.

Dimensions: -1:1 in x & y, 0:1 in Z.

Where Image space hidden surface algorithms work.

Screen Space (2D)

Coordinates 0:width, 0:height

Projections

We will look at several planar geometric 3D to 2D projection:

-Parallel Projections
Orthographic
Oblique

-Perspective

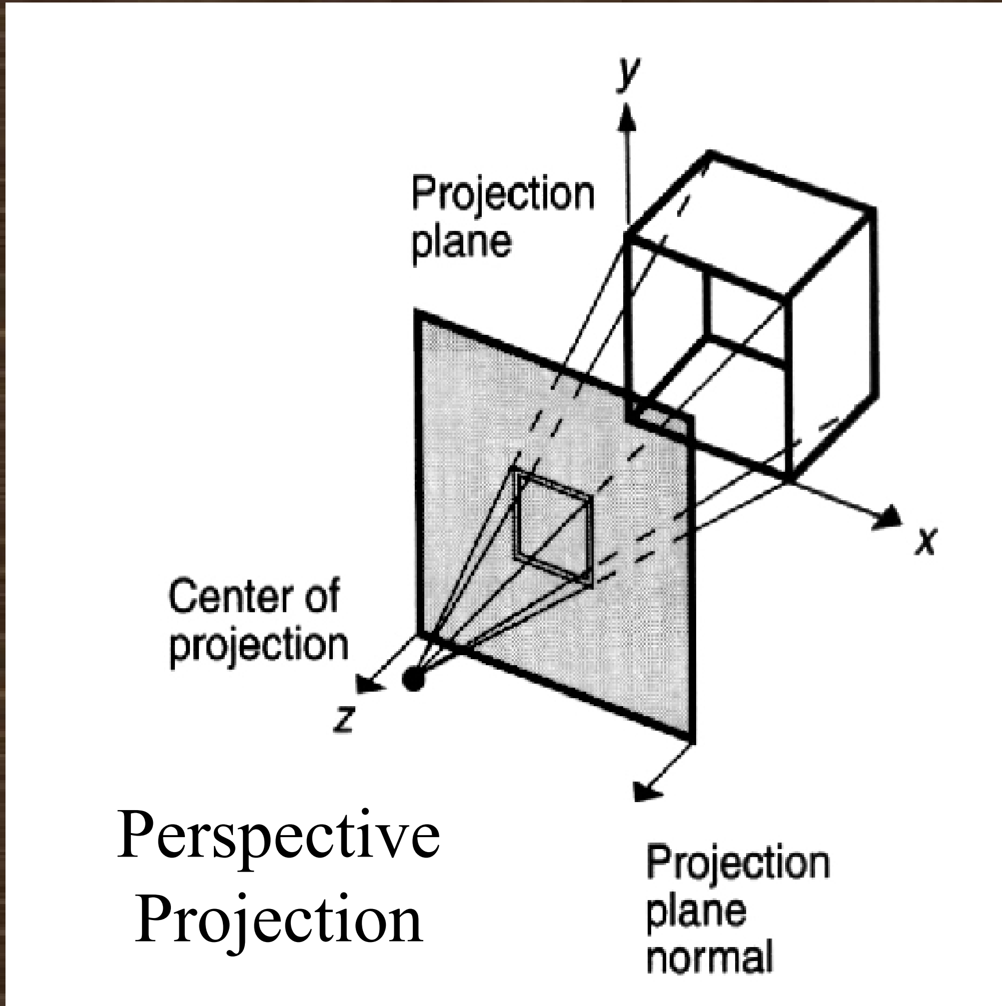
Projection of a 3D object is defined by straight projection rays (projectors) emanating from the center of projection (COP) passing through each point of the object and intersecting the projection plane.

Perspective Projections

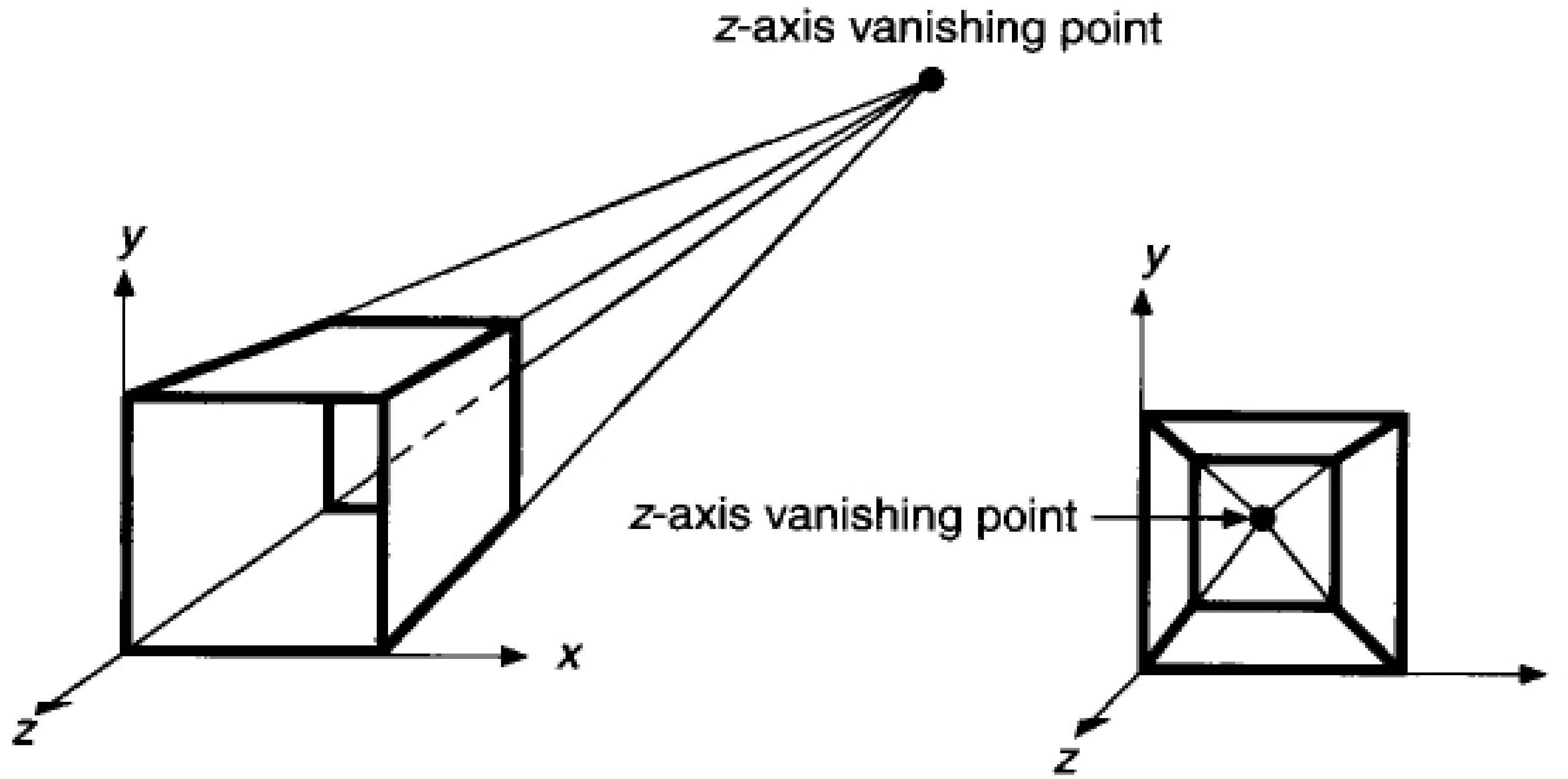
Distance from COP to projection plane is finite. The projectors are not parallel & we specify a center of projection.

Center of Projection is also called the Perspective Reference Point

COP = PRP

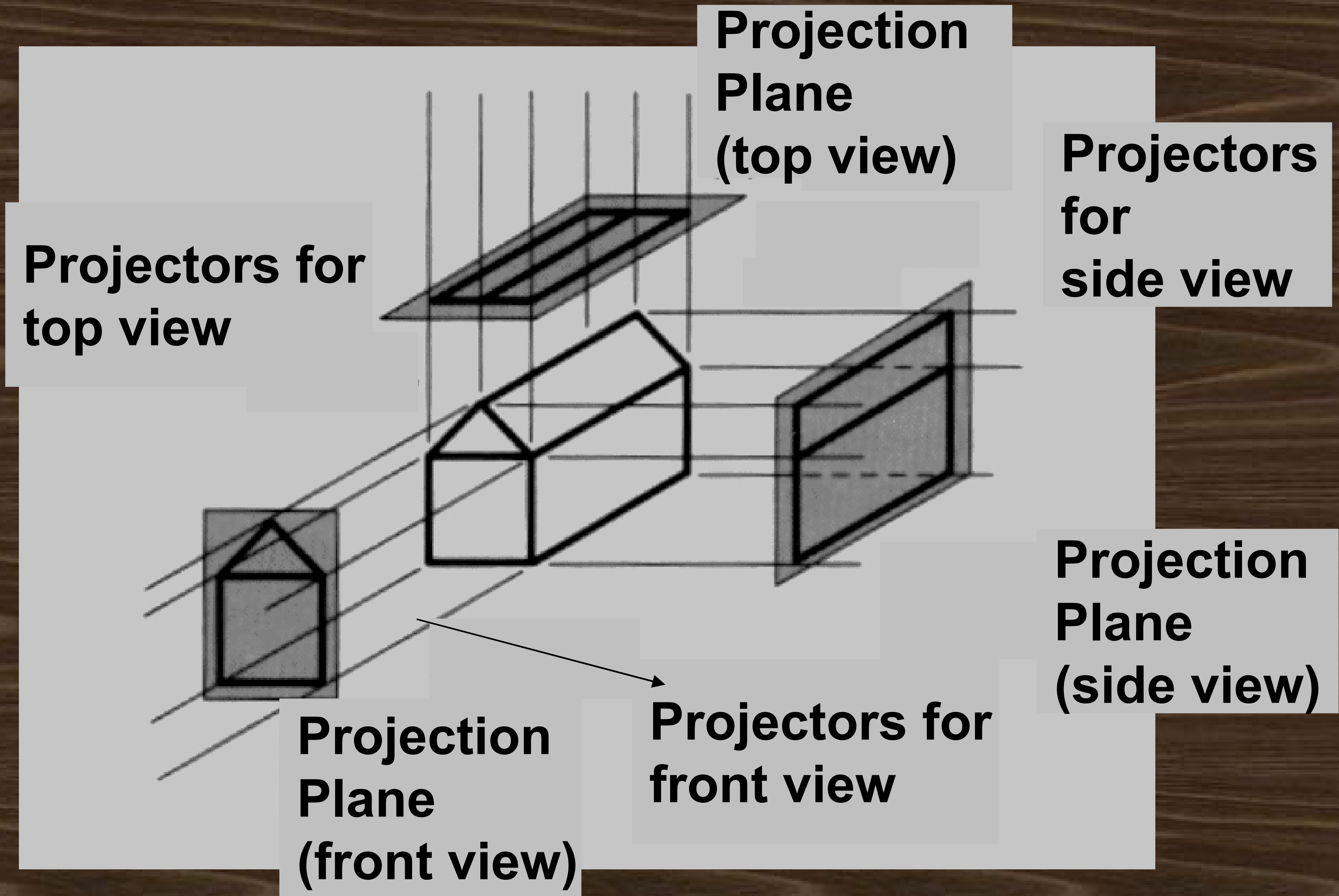


Perspective Projection



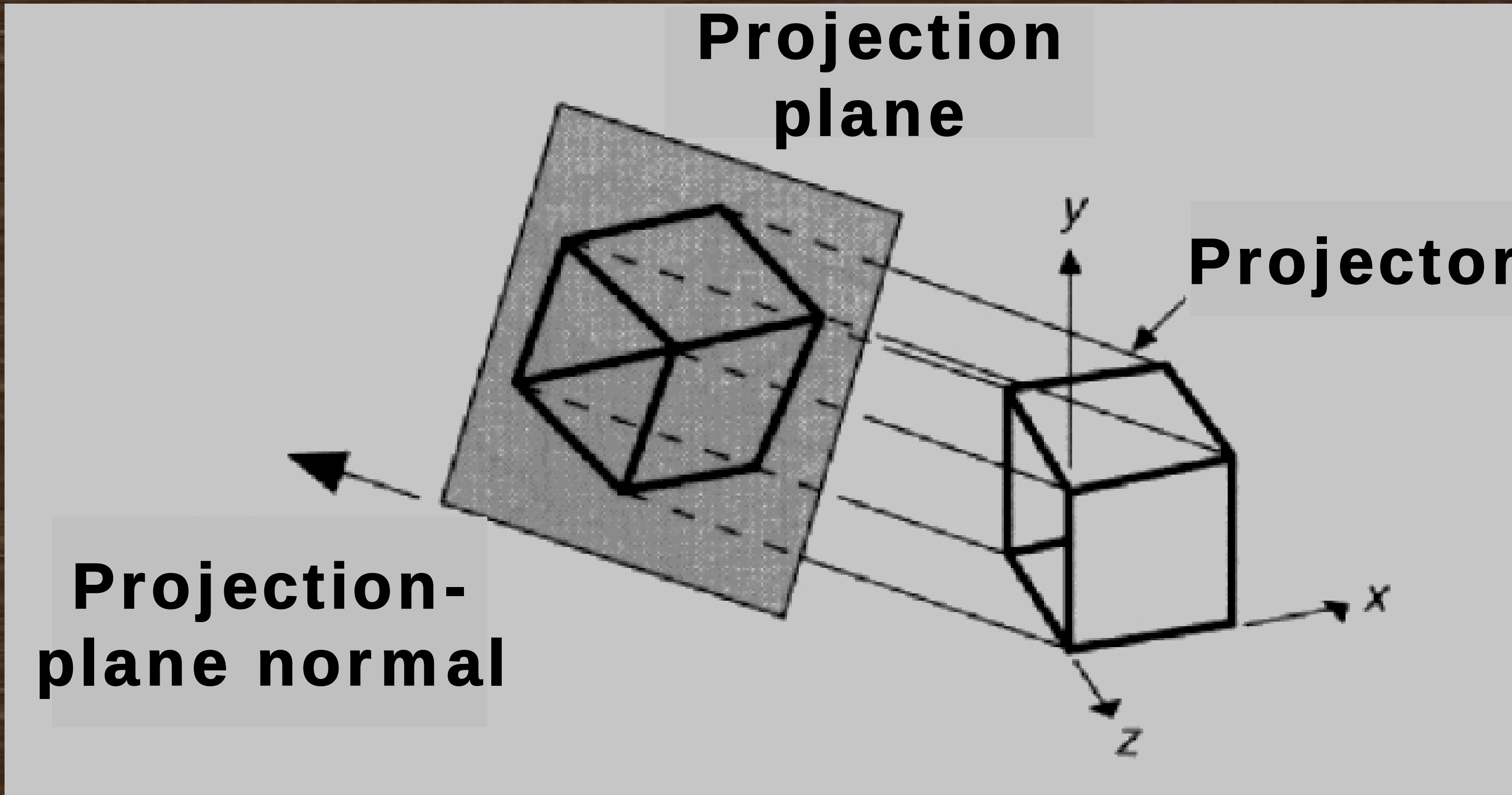
- **Perspective foreshortening:** the size of the perspective projection of the object varies inversely with the distance of the object from the center of projection.

- **Vanishing Point:** The perspective projections of any set of parallel lines that are not parallel to the projection plane converge to a vanishing point.

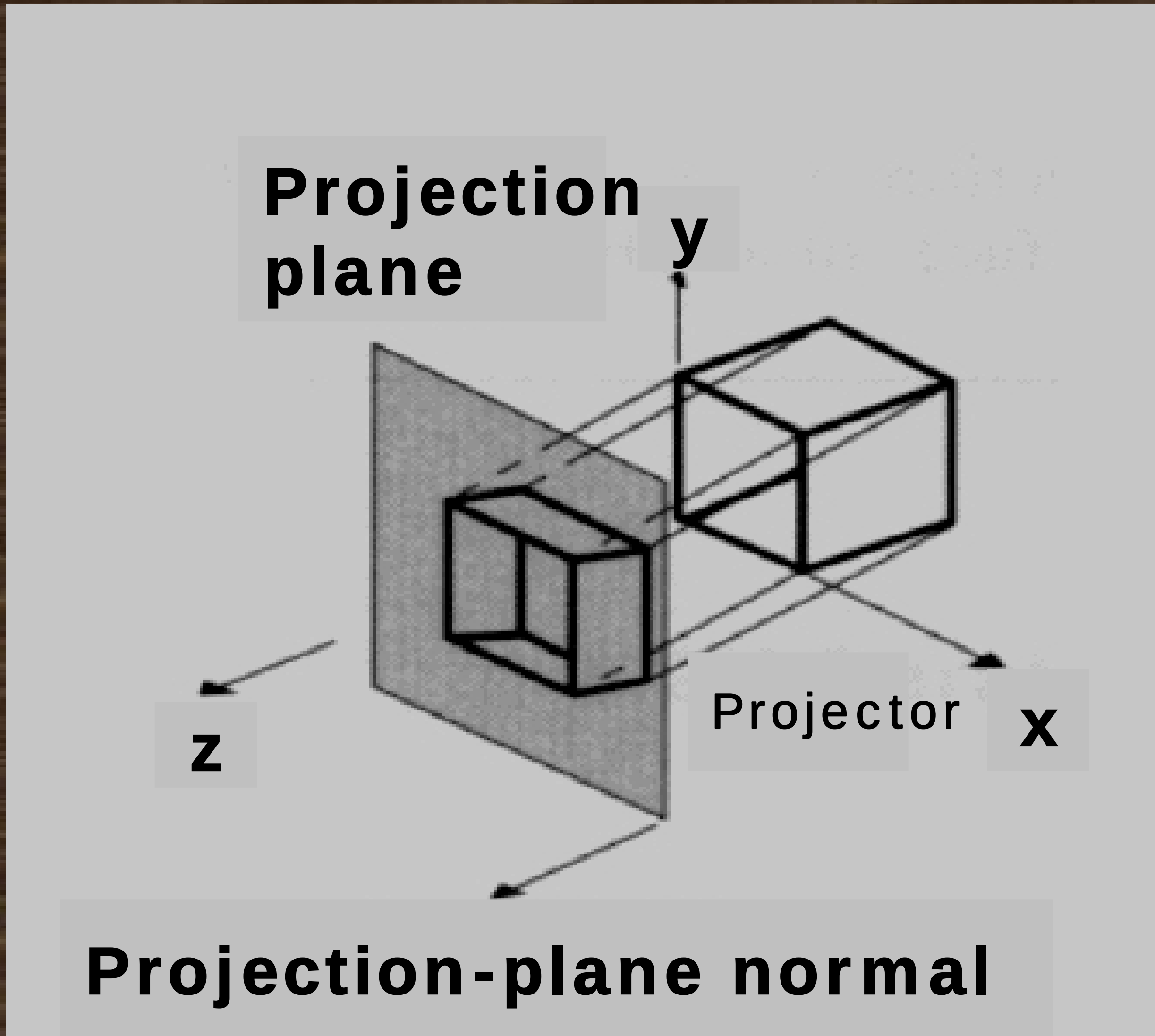


Example of Orthographic Projection

Example of Isometric Projection:

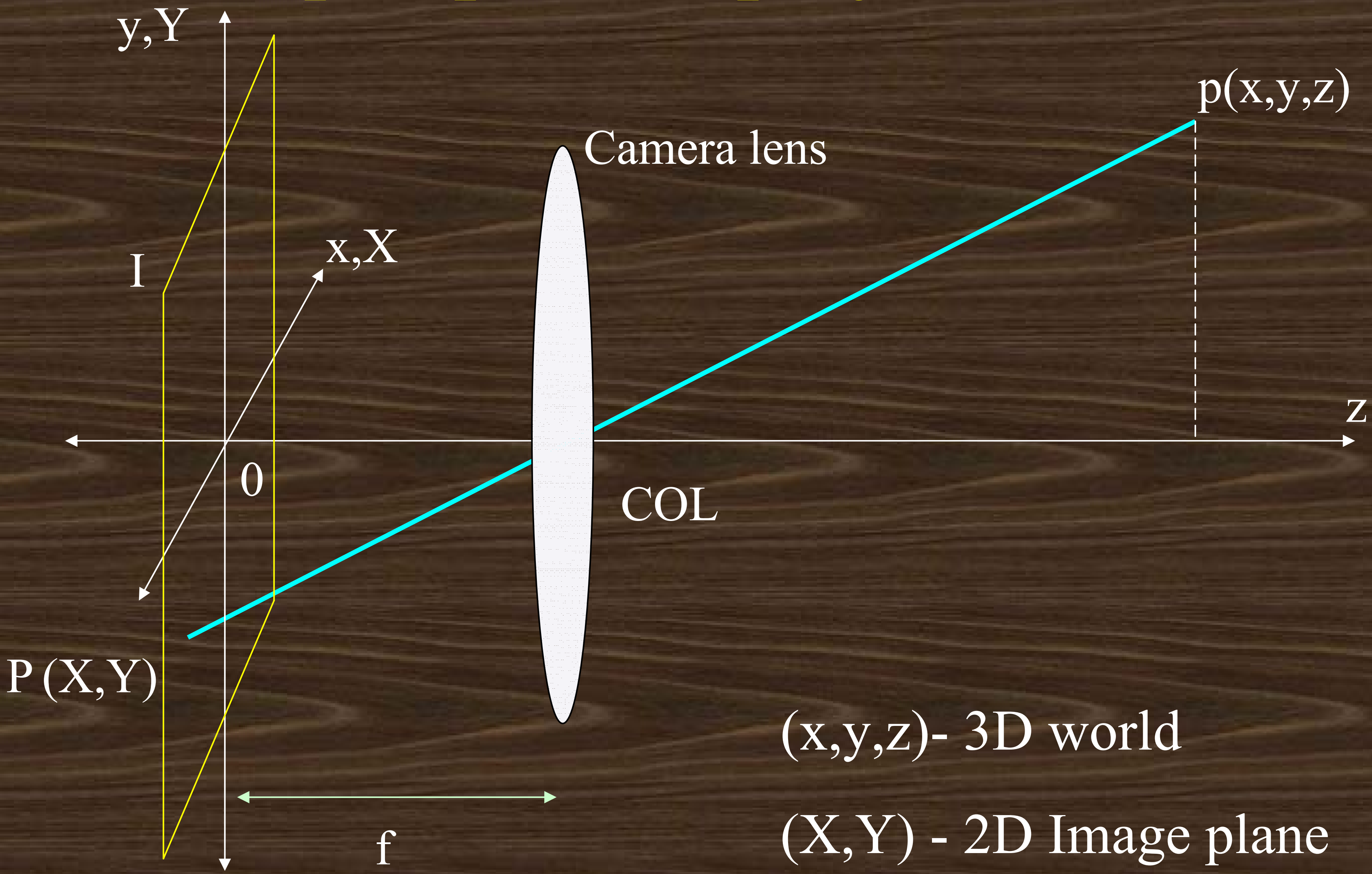


Example Oblique Projection



END OF BASICS

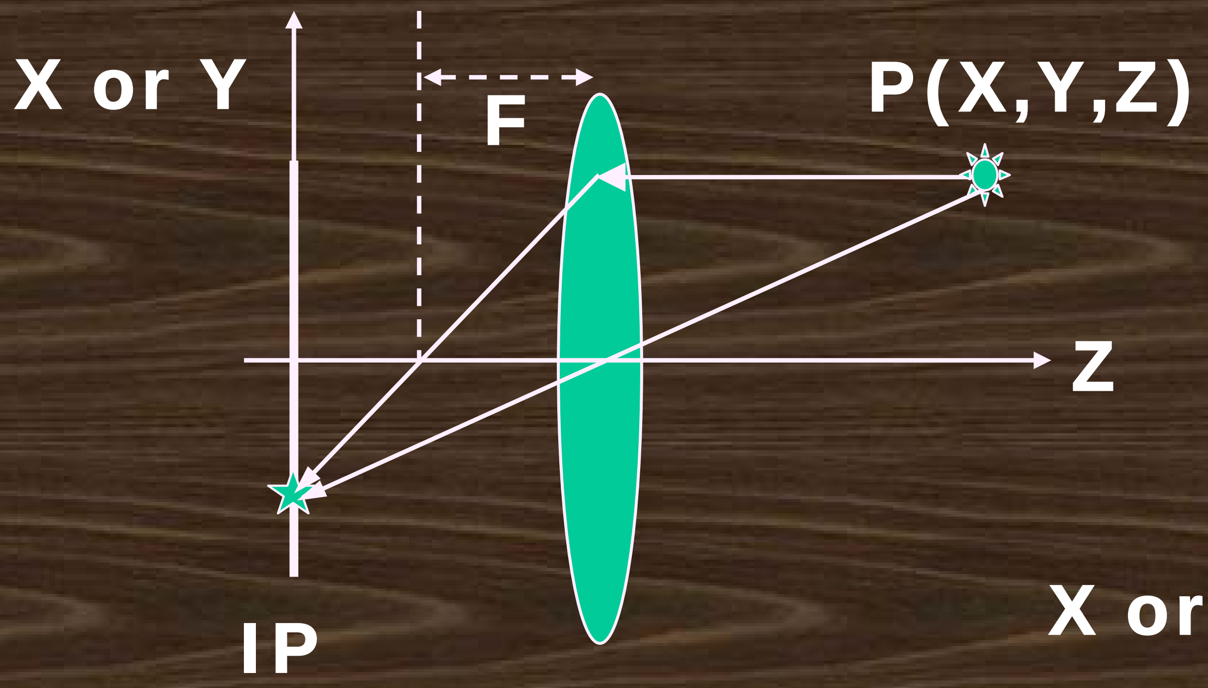
THE CAMERA MODEL: perspective projection

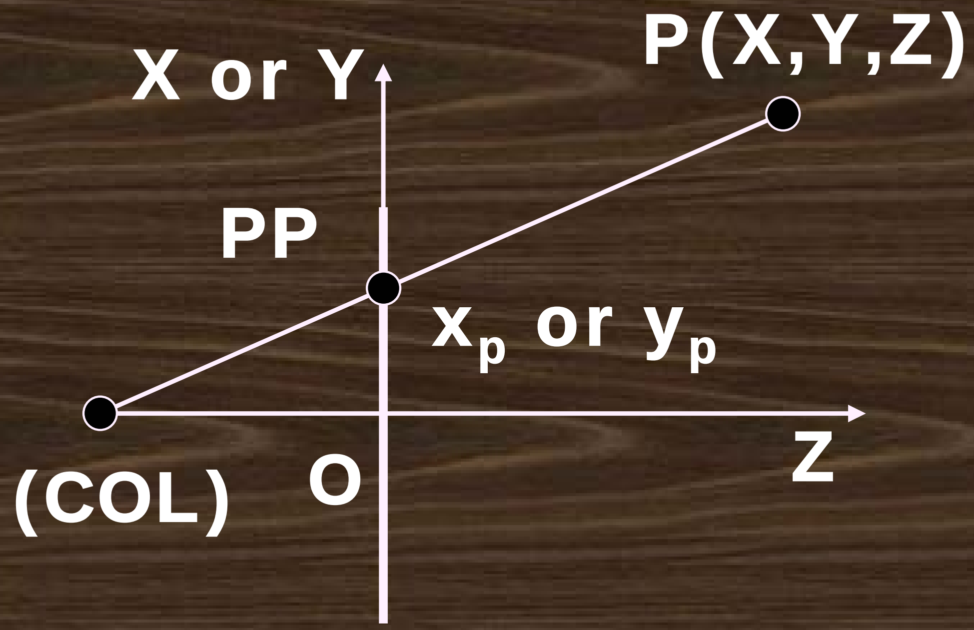


(x,y,z)- 3D world

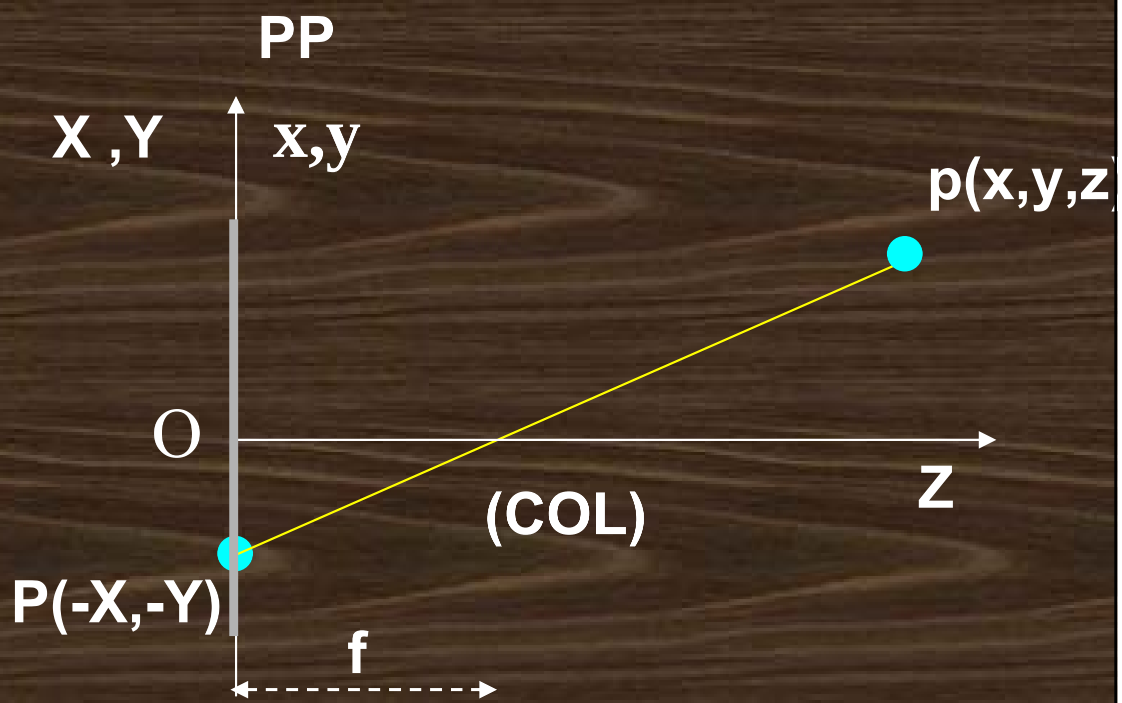
(X,Y) - 2D Image plane

Perspective Geometry and Camera Models





CASE 1



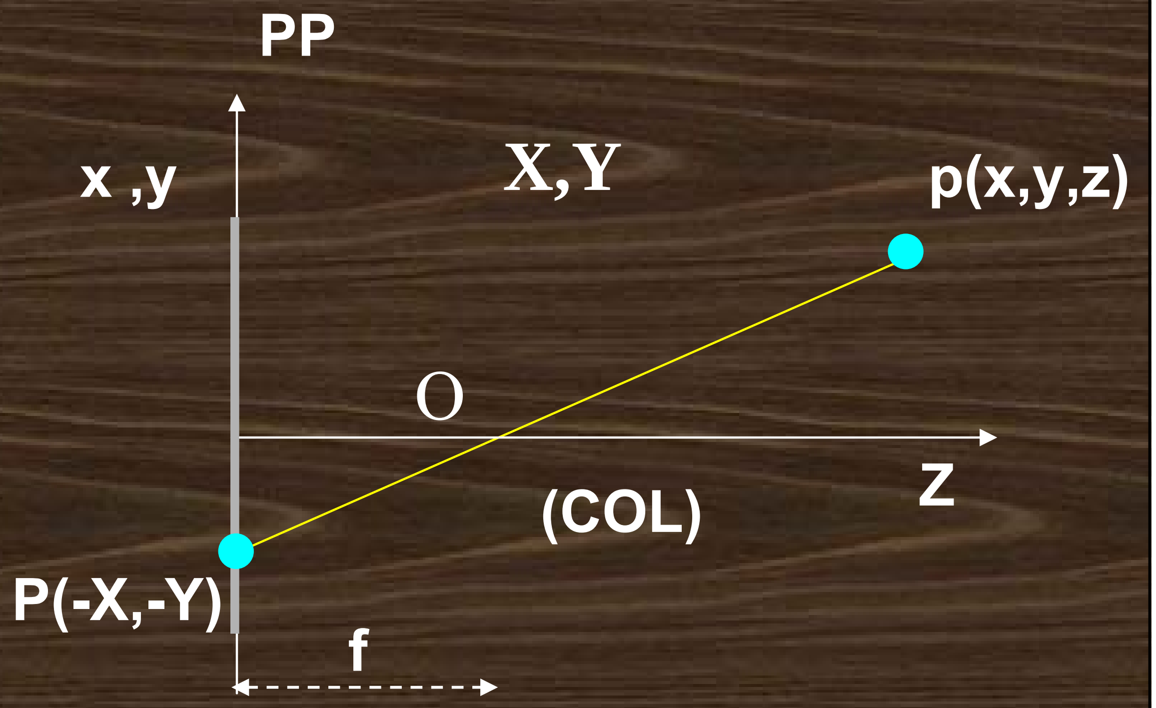
By similarity of triangles

$$\frac{X}{f} = \frac{-x}{z-f}, \qquad \frac{Y}{f} = \frac{-y}{z-f}$$

$$X = \frac{xf}{f-z}, \qquad Y = \frac{yf}{f\text{-}z}$$

$$X = \frac{x}{1 - z/f}, \quad Y = \frac{y}{1 - z/f}$$

- Image plane before the camera lens
- Origin of coordinate systems at the image plane
- Image plane at origin of coordinate system



CASE 2

By similarity of triangles

$$\frac{-X}{-f} = \frac{x}{z}, \qquad \frac{-Y}{-f} = \frac{y}{z}$$

$$X = \frac{xf}{z}, \qquad Y = \frac{yf}{z}$$

$$X = \frac{x}{z/f}, \quad Y = \frac{y}{z/f}$$

- Image plane before the camera lens
- Origin of coordinate systems at the camera lens
- Image plane at origin of coordinate system

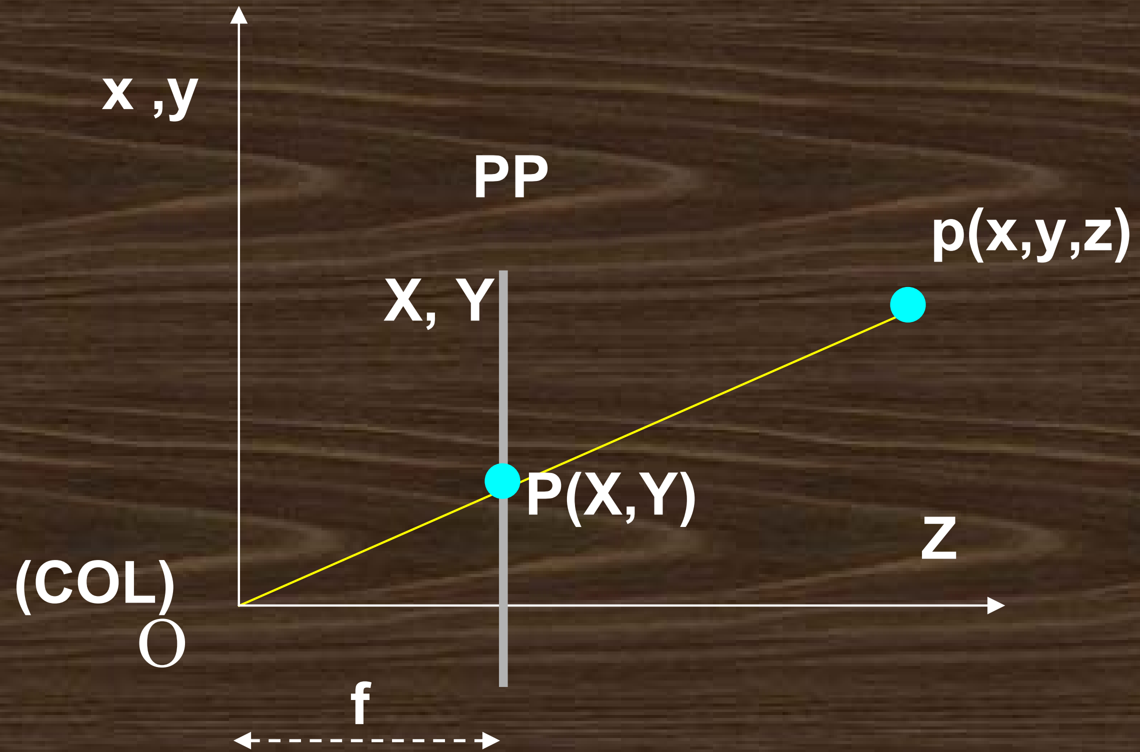
CASE 3

By similarity of triangles

$$\frac{X}{f} = \frac{x}{z}, \qquad \frac{Y}{f} = \frac{y}{z}$$

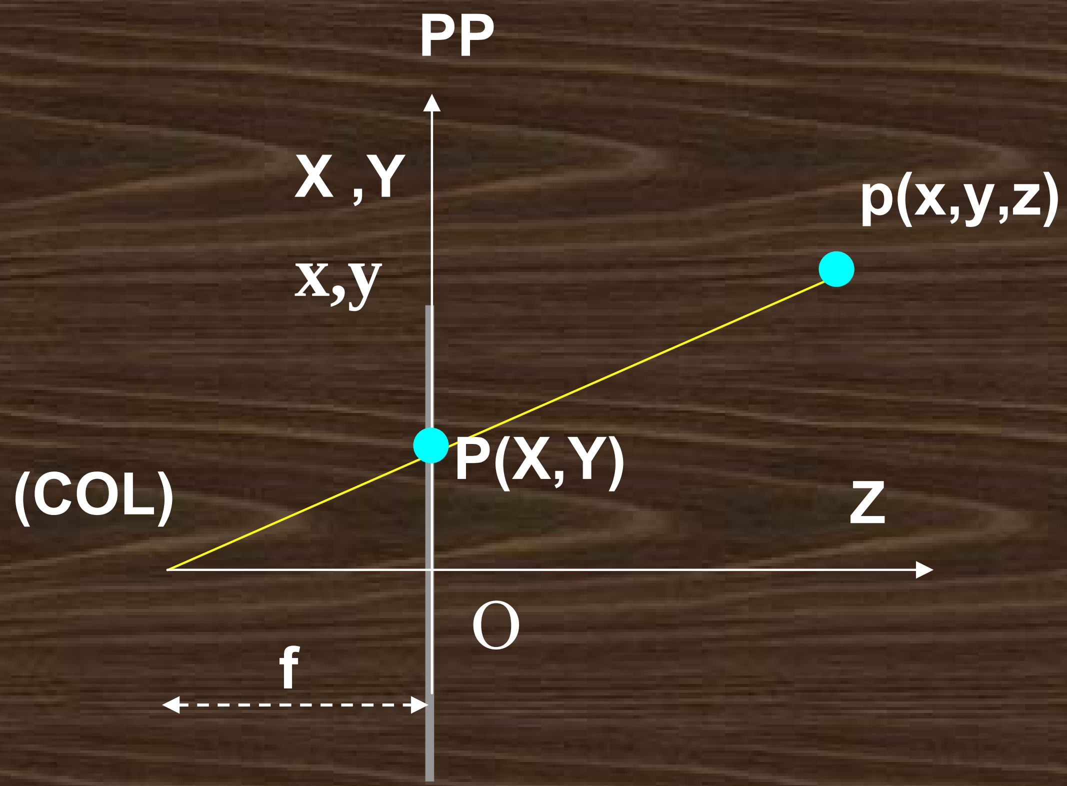
$$X = \frac{xf}{z}, \qquad Y = \frac{yf}{z}$$

$$X = \frac{x}{z/f}, \quad Y = \frac{y}{z/f}$$



- Image plane after the camera lens
- Origin of coordinate systems at the camera lens
- Focal length f

CASE 4



By similarity of triangles

$$\frac{X}{f} = \frac{x}{f+z}, \qquad \frac{Y}{f} = \frac{y}{f+z}$$

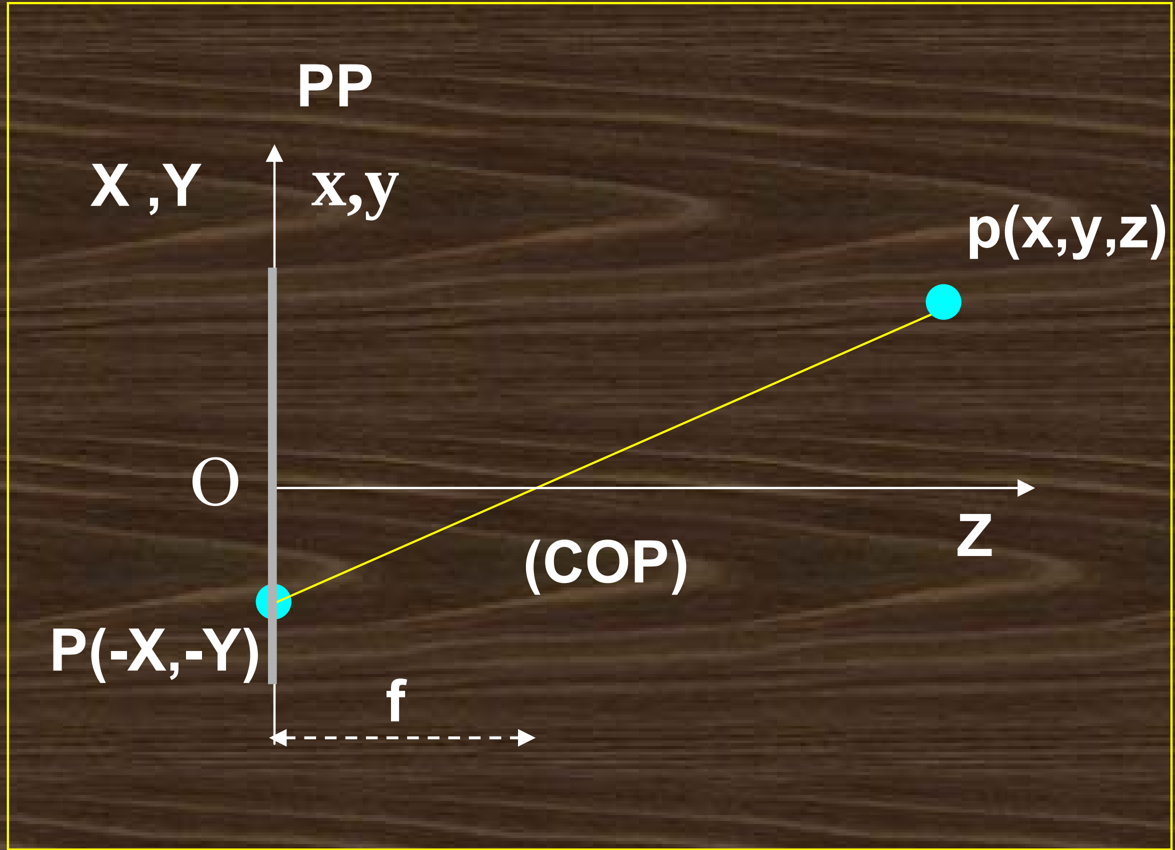
$$X = \frac{xf}{f+z}, \qquad Y = \frac{yf}{f+z}$$

$$X = \frac{x}{1+ z/f}, \qquad Y = \frac{y}{1+ z/f}$$

- Image plane after the camera lens
- Origin of coordinate system not at COP
- Image plane origin coincides with 3D world origin

Consider the first case ….

- Note that the equations are non-linear
- We can develop a matrix formulation of the equations given below



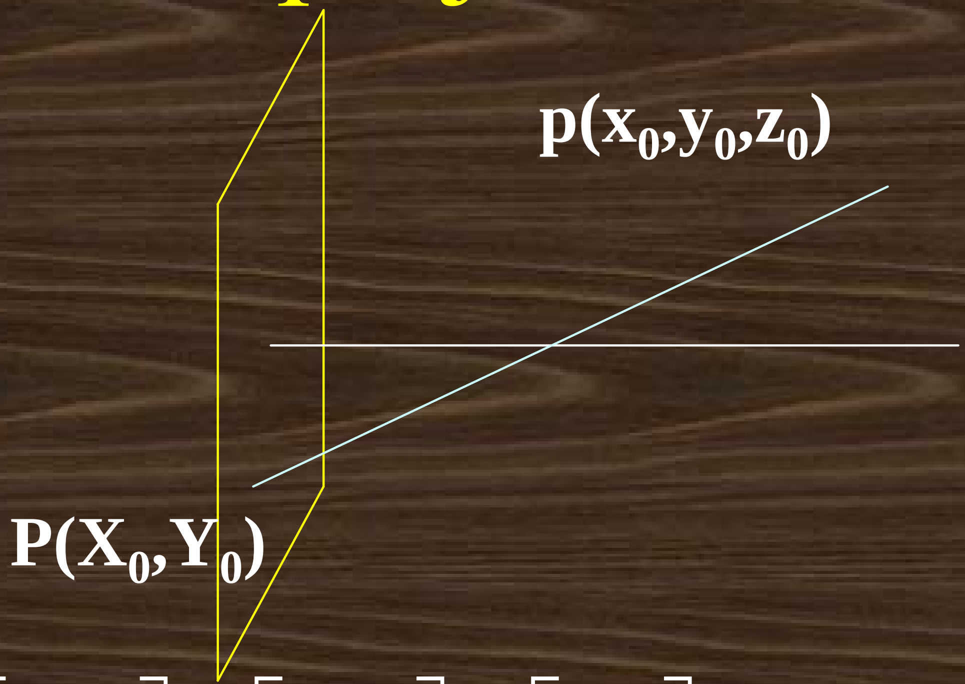
$$X = \frac{x}{1 - z/f}, \quad Y = \frac{y}{1 - z/f}$$

$$\begin{bmatrix} X \\ Y \\ Z \\ k' \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1/f & 1 \end{bmatrix} \begin{bmatrix} kx \\ ky \\ kz \\ k \end{bmatrix}$$

(Z is not important and is eliminated)

Inverse perspective projection

$$P^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix}$$



$$w_h = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} kX_0 \\ kY_0 \\ 0 \\ k \end{bmatrix} = \begin{bmatrix} kX_0 \\ kY_0 \\ 0 \\ k \end{bmatrix} = \begin{bmatrix} X_0 \\ Y_0 \\ 0 \\ 1 \end{bmatrix}$$

Hence no 3D information can be retrieved with the inverse transformation

So we introduce the dummy variable i.e. the depth Z

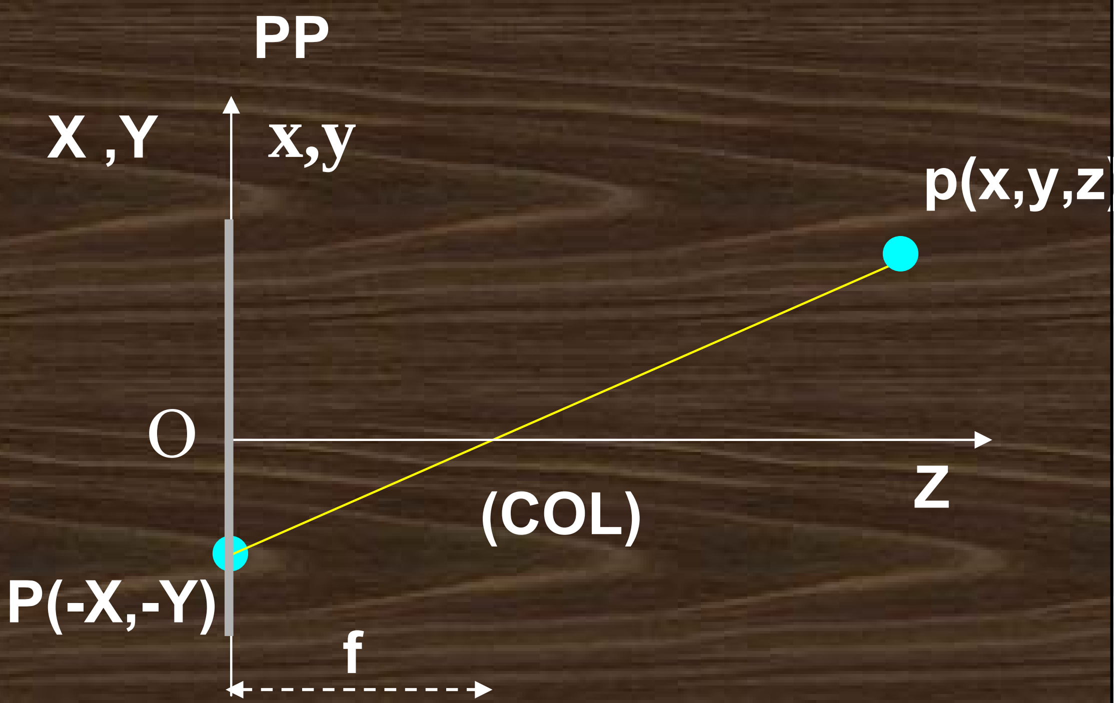
Let the image point be represented as: $[kX_0 \ \ kY_0 \ \ kZ \ \ k]^T$

$$w_h = \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 1 \end{bmatrix} \begin{bmatrix} kX_0 \\ kY_0 \\ kZ \\ k \end{bmatrix} = \begin{bmatrix} kX_0 \\ kY_0 \\ kZ \\ k\,Z/f + k \end{bmatrix} = \begin{bmatrix} fX_0 / (f+Z) \\ fY_0 / (f+Z) \\ fZ / (f+Z) \\ 1 \end{bmatrix}$$

$$z_0 = \frac{fZ}{f+Z} \Rightarrow Z = \frac{fz_0}{f-z_0} \Rightarrow \frac{f}{f+Z} = \frac{z_0}{Z} = \frac{f-z_0}{f}$$

$$\Rightarrow \quad x_0 = \frac{X_0}{f}(f-z_0), \qquad y_0 = \frac{Y_0}{f}(f-z_0)$$

CASE 1



Forward: 3D to 2D

$$\frac{X}{f} = \frac{-x}{z-f}, \qquad \frac{Y}{f} = \frac{-y}{z-f}$$

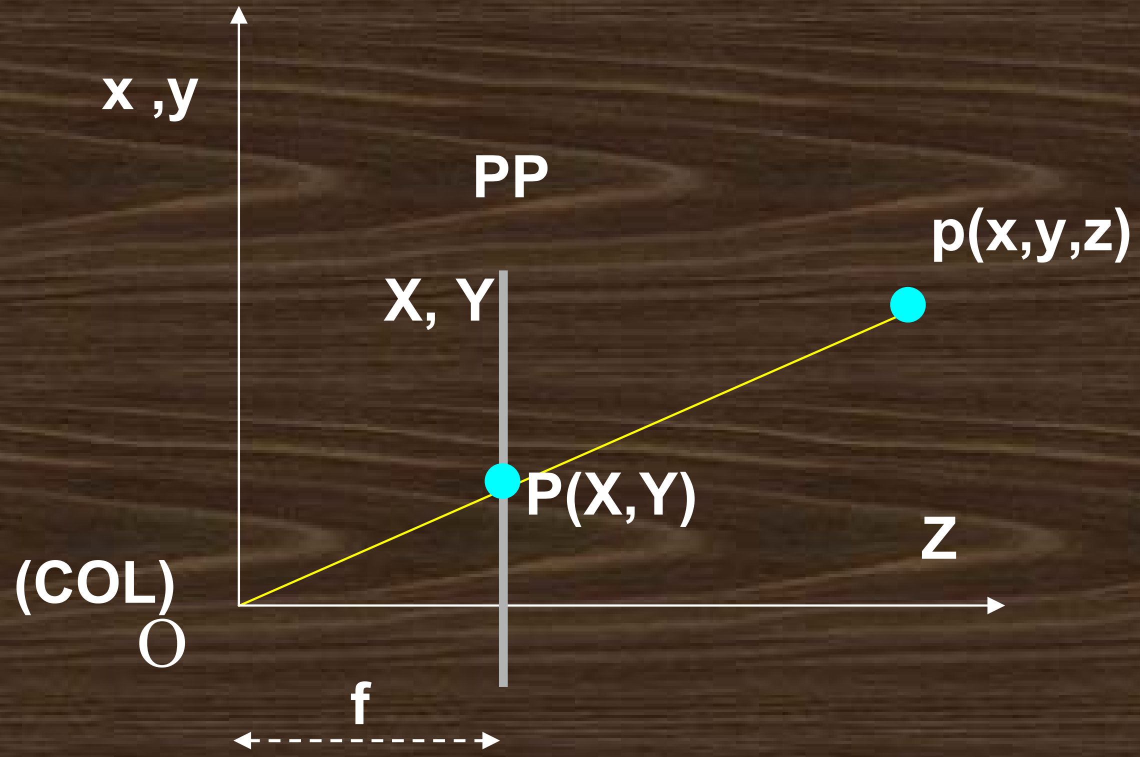
$$X = \frac{xf}{f-z}, \qquad Y = \frac{yf}{f\text{-}z}$$

$$X = \frac{x}{1 - z/f}, \qquad Y = \frac{y}{1 - z/f}$$

Inverse: 2D to 3D

$$x_0 = \frac{X_0}{f}(f - z_0), \qquad y_0 = \frac{Y_0}{f}(f - z_0)$$

CASE 3



Forward: 3D to 2D

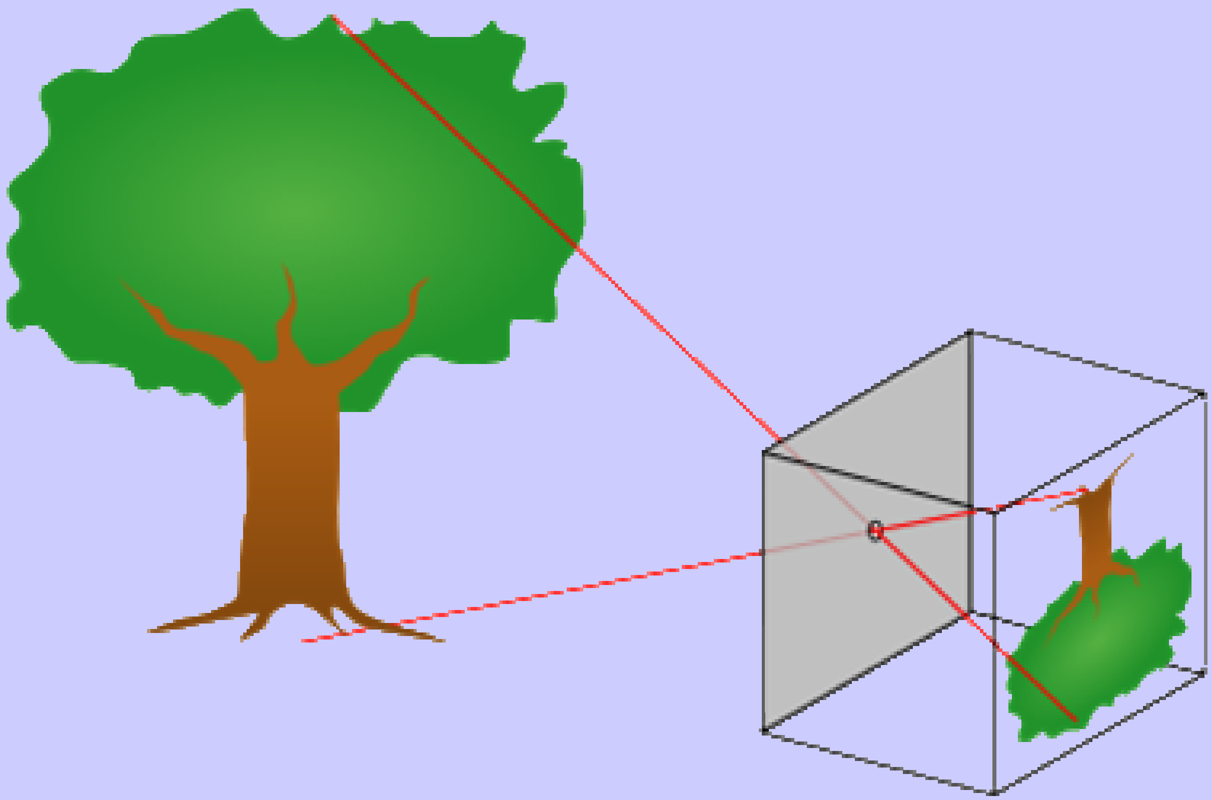
$$\frac{X}{f} = \frac{x}{z}, \qquad \frac{Y}{f} = \frac{y}{z}$$

$$X = \frac{xf}{z}, \qquad Y = \frac{yf}{z}$$

$$X = \frac{x}{z/f}, \quad Y = \frac{y}{z/f}$$

Inverse: 2D to 3D

$$x_0 = \frac{z_0.X_0}{f}, \qquad y_0 = \frac{z_0.Y_0}{f}$$

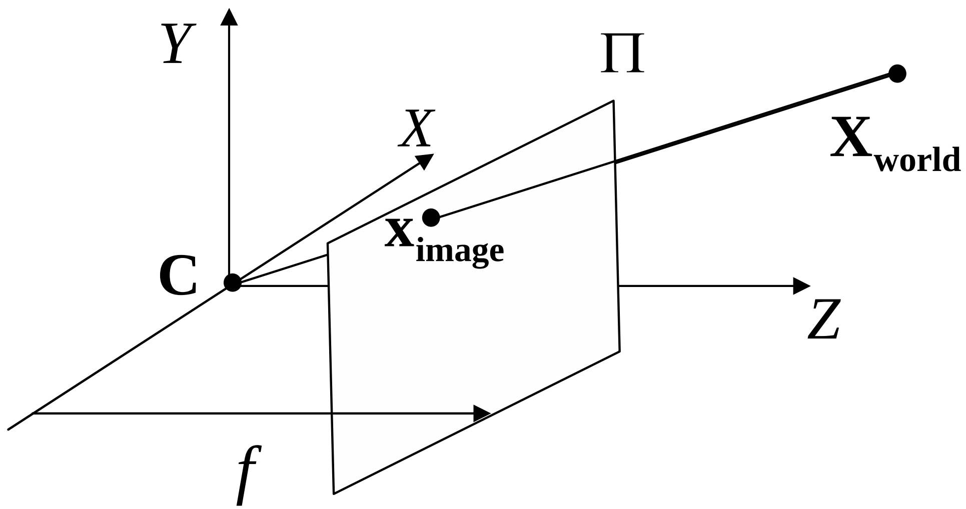


Pinhole Camera schematic diagram

Camera Image formulation

- Action of eye is simulated by an abstract camera model (pinhole camera model)
- 3D real world is captured on the image plane. Image is projection of 3D object on a 2D plane.

$$F:(X_w, Y_w, Z_w) \rightarrow (x_i, y_i)$$



$$\mathbf{C} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{pmatrix} \sim \begin{pmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$$

$$\mathbf{X}_{\text{world}} = (X_w, Y_w, Z_w)$$

$$\mathbf{x}_{\text{image}} = \left(f\frac{X_w}{Z_w}, f\frac{Y_w}{Z_w}\right)$$

Camera Geometry

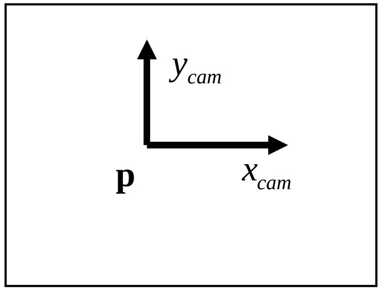
- Camera can be considered as a projection matrix, $\mathbf{x} = \mathrm{P}_{3*4}\mathbf{X}$
  - A pinhole camera has the projection matrix as

$$P = diag(f, f, 1)[I \quad | \quad 0]$$

  - Principal point offset

$$(X, Y, Z)^{\mathrm{T}} \rightarrow \left(fX / Z + p_x, fY / Z + p_y\right)^{\mathrm{T}}$$

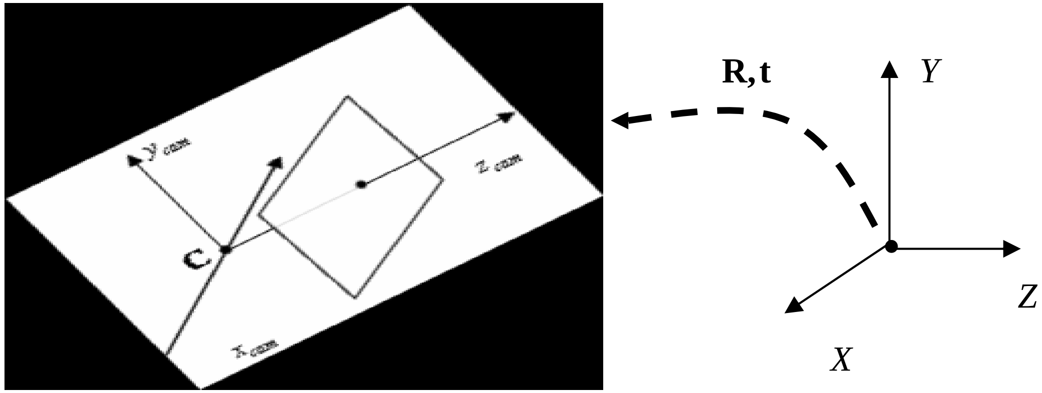
$$K = \begin{bmatrix} f & 0 & p_x & 0 \\ 0 & f & p_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \qquad \mathbf{x} = \mathrm{K}[\mathrm{I} \quad | \quad \mathbf{0}]\mathbf{X}$$



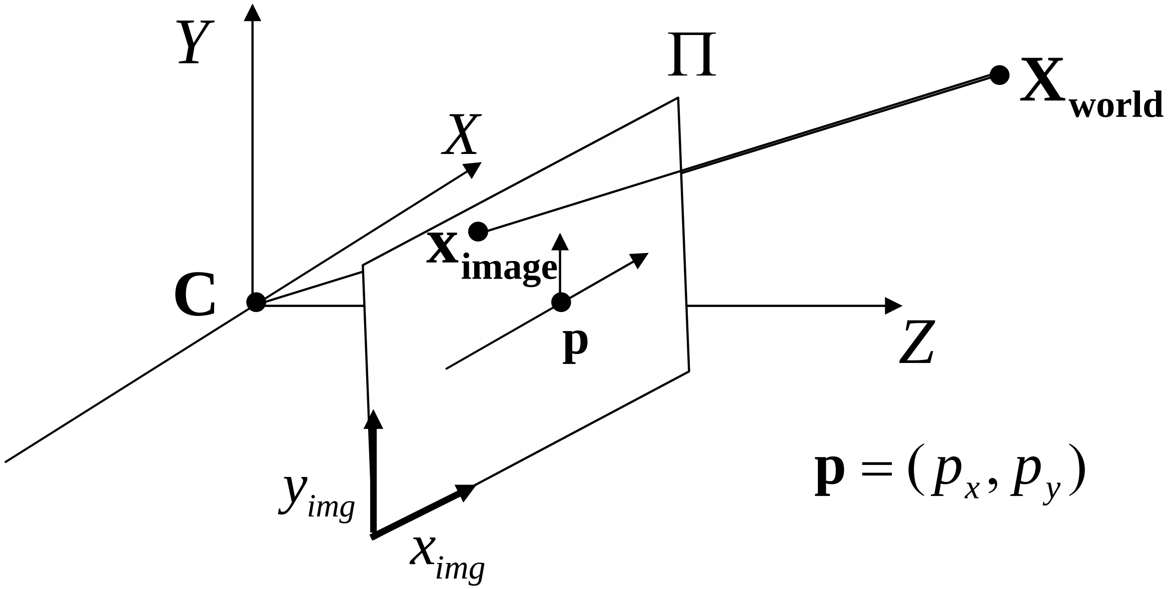
$\mathbf{p} = (p_x, p_y)$

  - Camera with rotation and translation

$$\mathbf{x} = \mathrm{K}[\mathrm{R} \mid \mathbf{t}]\mathbf{X}$$



Camera Geometry



$$\mathbf{p} = (p_x, p_y)$$

Camera internal parameters

$$K = \begin{bmatrix} \alpha_x & s & p_x \\ & \alpha_y & p_y \\ & & 1 \end{bmatrix}$$

α_x Scale factor in x- coordinate direction

α_y Scale factor in y- coordinate direction

s Camera skew

$\frac{\alpha_x}{\alpha_y}$ Aspect ratio

Camera matrix, $\boxed{P = K[R \mid \mathbf{t}]}$

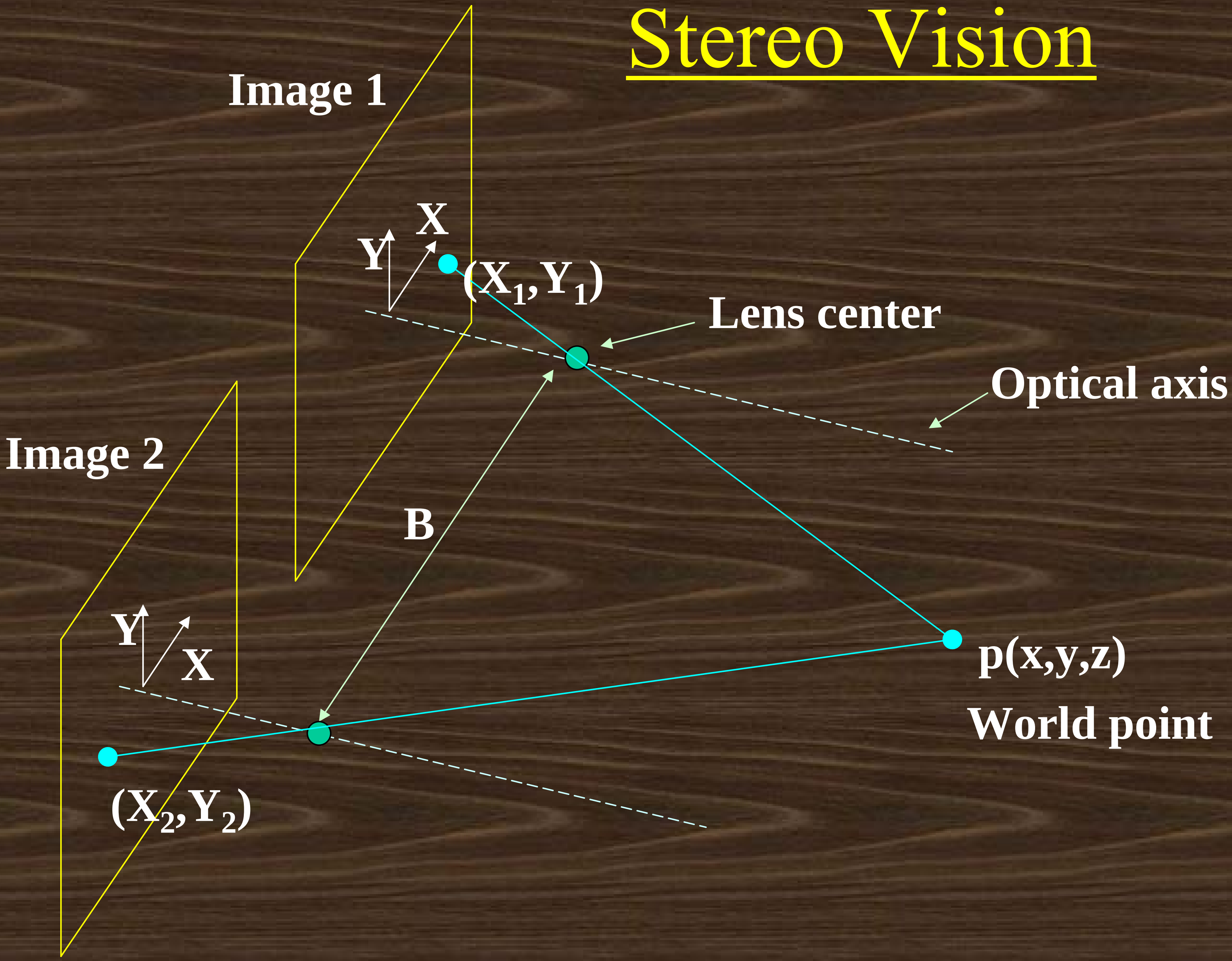
R Rotation

$\mathbf{t}$ Translation vector

Observations about Perspective projection

- 3D scene to image plane is a one to one transformation (unique correspondence)
- For every image point no unique world coordinate can be found
- So depth information cannot be retrieved using a single image ? What to do?
- Would two (2) images of the same object (from different viewing angles) help?
- Termed - Stereo Vision

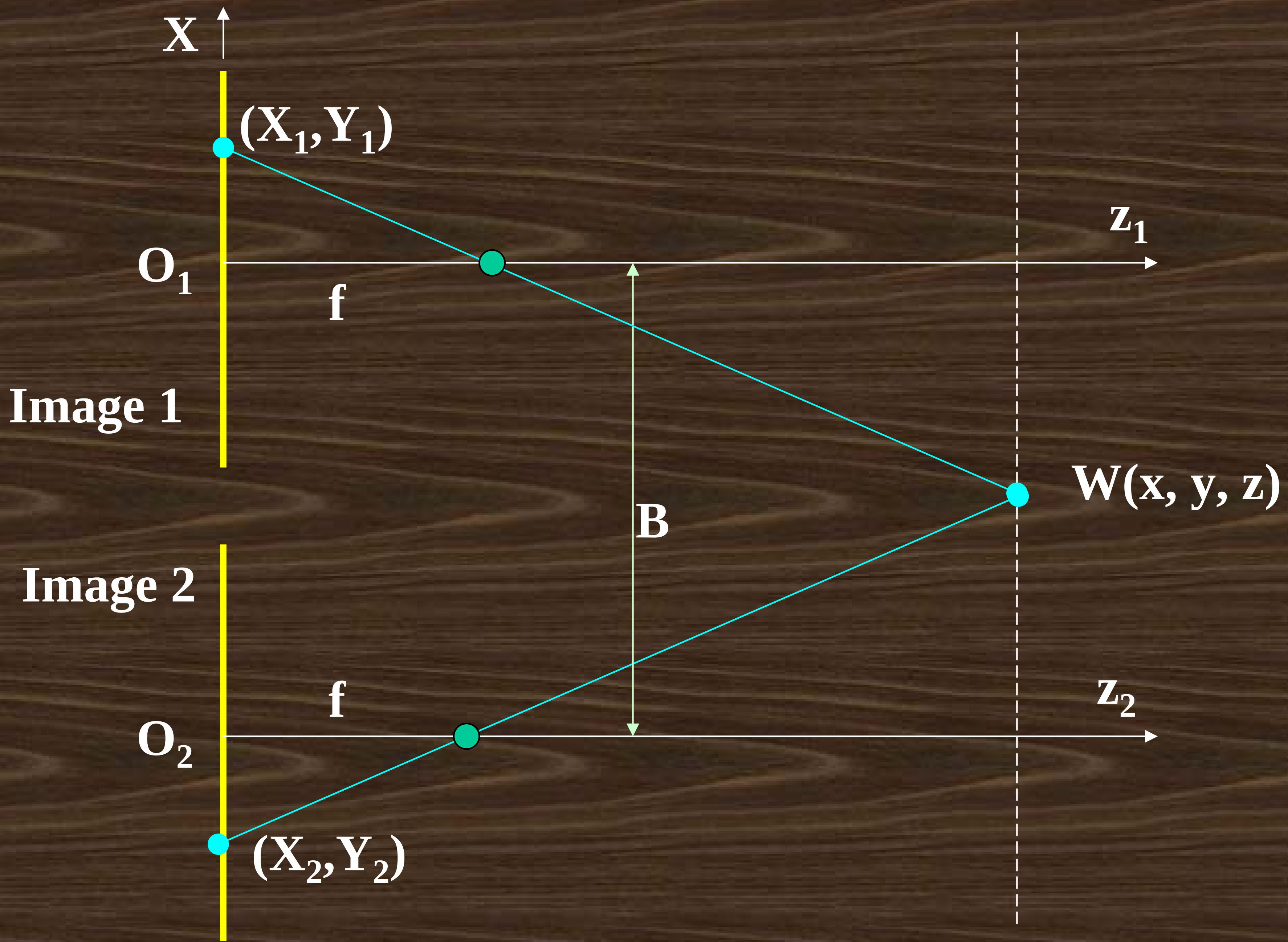
Stereo Vision



Stereo Vision (2)

- **Stereo imaging involves obtaining two separate image views of an object (in this discussion the world point)**
- **The distance between the centers of the two lenses is called the baseline width.**
- **The projection of the world point on the two image planes is (X_1, Y_1) and (X_2, Y_2)**
- **The assumption is that the cameras are identical**
- **The coordinate system of both cameras are perfectly aligned differing only in the x-coordinate location of the origin.**
- **The world coordinate system is also bought into the coincidence with one of the image X, Y planes (say image plane 1) . So y, z coordinates are same for both the camera coordinate systems.**

Top view of the stereo imaging system with origin at center of first imaging plane.



First bringing the first camera into coincidence with the world coordinate system and then using the second camera coordinate system and directly applying the formula we get:

$$x_1 = \frac{X_1}{f}(f - z_1), \quad x_2 = \frac{X_2}{f}(f - z_2)$$

Because the separation between the two cameras is B

$$x_2 = x_1 + B, \quad z_1 = z_2 = z(?) \quad /*\,\text{Solve it now}\,*/$$

$$x_1 = \frac{X_1}{f}(f - z), \quad x_1 + B = \frac{X_2}{f}(f - z)$$

$$B = \frac{(X_2 - X_1)}{f}(f - z), \quad z = f - \frac{fB}{(X_2 - X_1)}$$

- The equation above gives the depth directly from the coordinate of the two points
- The quantity given below is called the disparity

$$D = (X_2 - X_1) = \frac{fB}{(f - z)}$$

- The most difficult task is to find out the two corresponding points in different images of the same scene – **the correspondence problem**.
- Once the correspondence problem is solved – (non-analytical), we get D. Then obtain depth using:

$$\mathbf{z = f - \frac{fB}{(X_2 - X_1)} = f\,[1 - B/D]}$$

Alternate Model – Case III

$$\frac{X}{f} = \frac{x}{z}, \quad \frac{Y}{f} = \frac{y}{z}$$

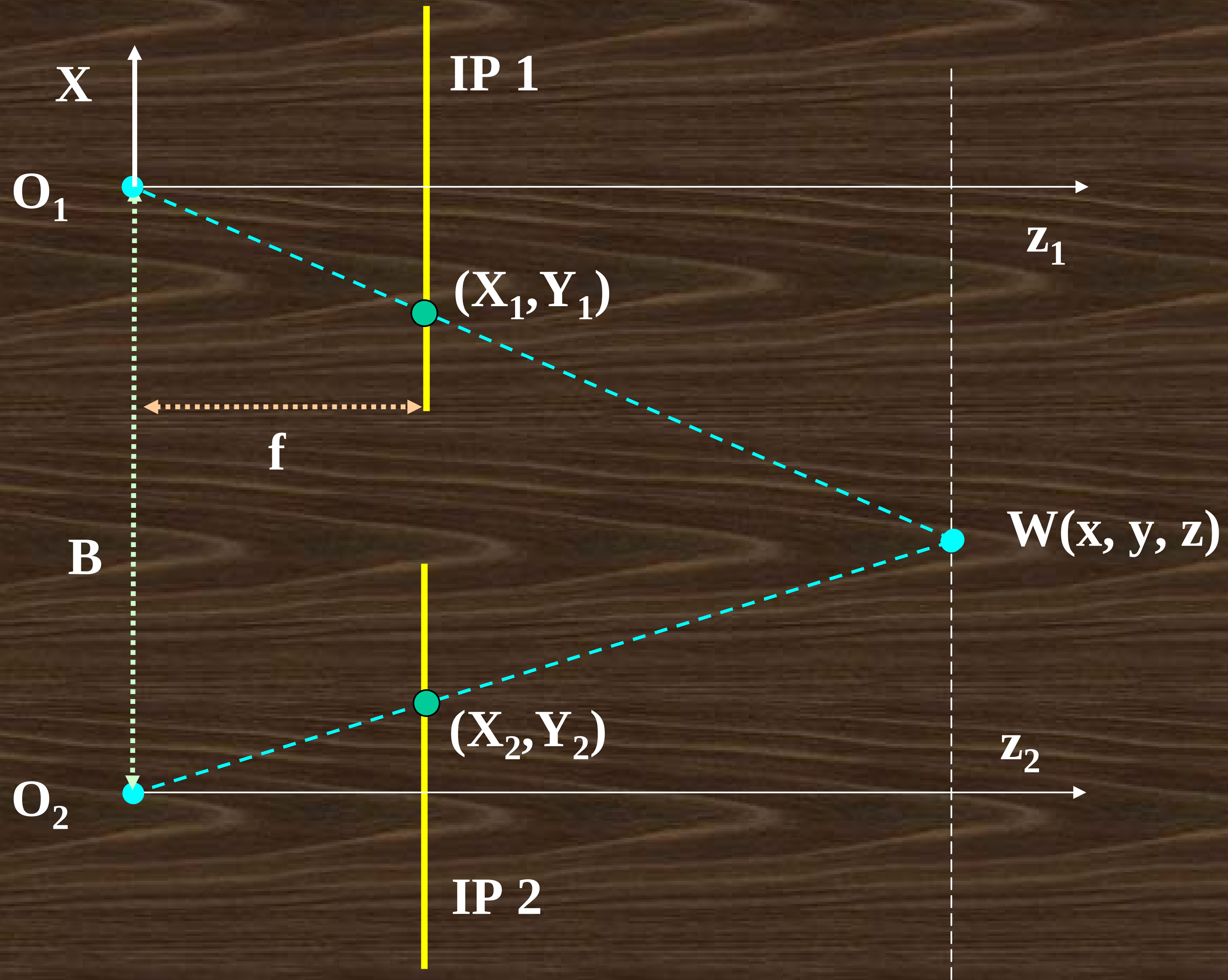
$$x = \frac{Xz}{f}, \quad y = \frac{Yz}{f}$$

$$x_2 = x_1 - B, \quad y_1 = y_2 = y; \quad z_1 = z_2 = z(?).$$

$$x_1 = \frac{X_1 z}{f}, \quad x_2 = x_1 - B = \frac{X_2 z}{f}$$

$$B = \frac{(X_1 - X_2)z}{f}, \quad z = \frac{fB}{(X_1 - X_2)} = B.f \big/ D$$

Top view of the stereo imaging system with origin at center of first camera lens.



Compare the two solutions

$$z = f - \frac{fB}{(X_2 - X_1)} = f[1 - B/D]$$

$$D = (X_2 - X_1) = \frac{fB}{(f - z)}$$

$$z = \frac{fB}{(X_1 - X_2)} = B.f/D$$

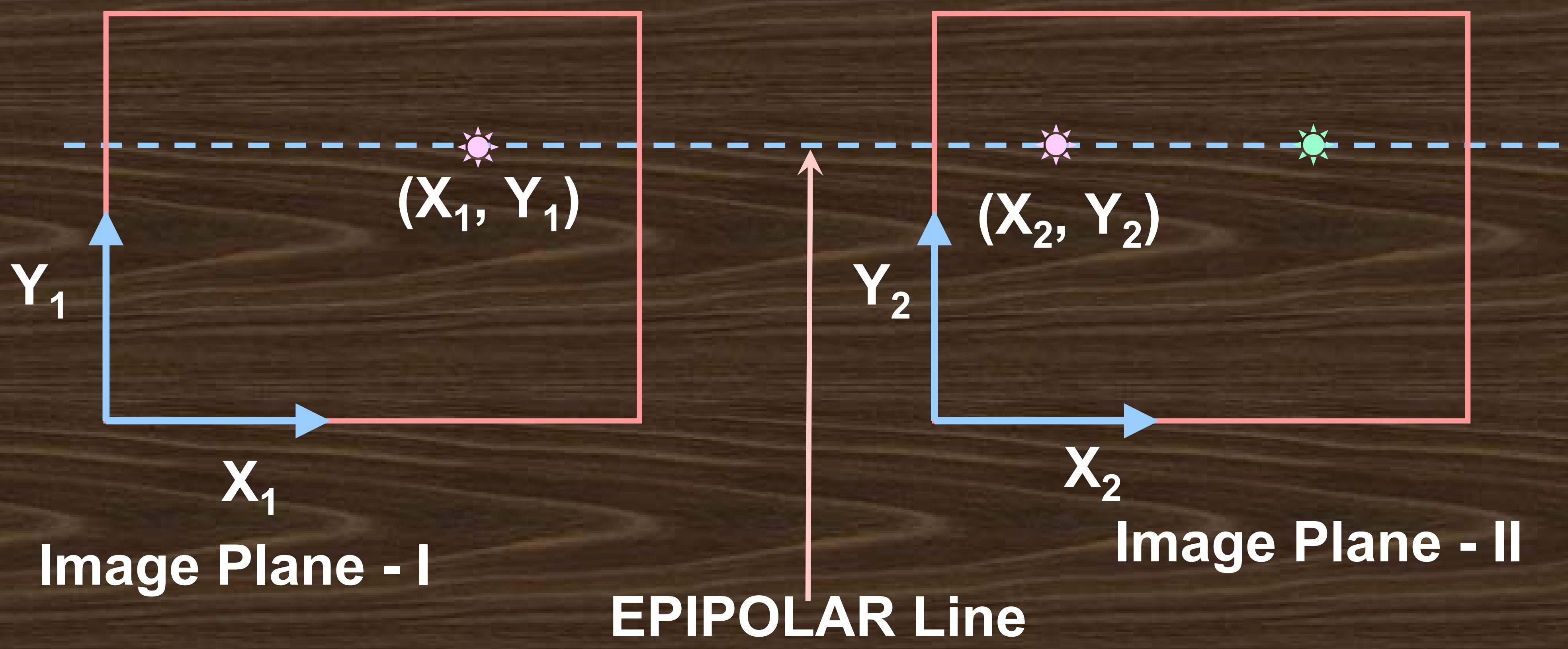
$$D = (X_1 - X_2) = \frac{fB}{z}$$

What do you think of D ?

The Correspondence Problem

$$z = {}^{B.f}\!/_{D} \qquad D = (X_1 - X_2) = \frac{fB}{z} \qquad Y_1 = Y_2$$

$$If \quad D > 0; \quad then \;\; X_2 < X_1$$



Error in Depth Estimation

$$z = {B.f}/{D} \qquad {\delta(z)}/{\delta D} = -{B.f}/{D^2}$$

Expressing in terms of depth (z), we have:

$${\delta(z)}/{\delta D} = -\frac{B.f}{D^2} = -\frac{z}{D} = -\frac{z^2}{B.f}$$

What is the maximum value of depth (z), you can measure using a stereo setup ?

$$z_{\max} = B.f$$

Even if correspondence is solved correctly, the computation of D may have an error, with an upper bound of 0.5; i.e. $(\delta D)_{max} = 0.5$.

That may cause an error of: $$\delta(z) = -\frac{z^2}{2B.f}$$

Larger baseline width and Focal length (of the camera) reduces the error and increases the maximum value of depth that may be estimated.

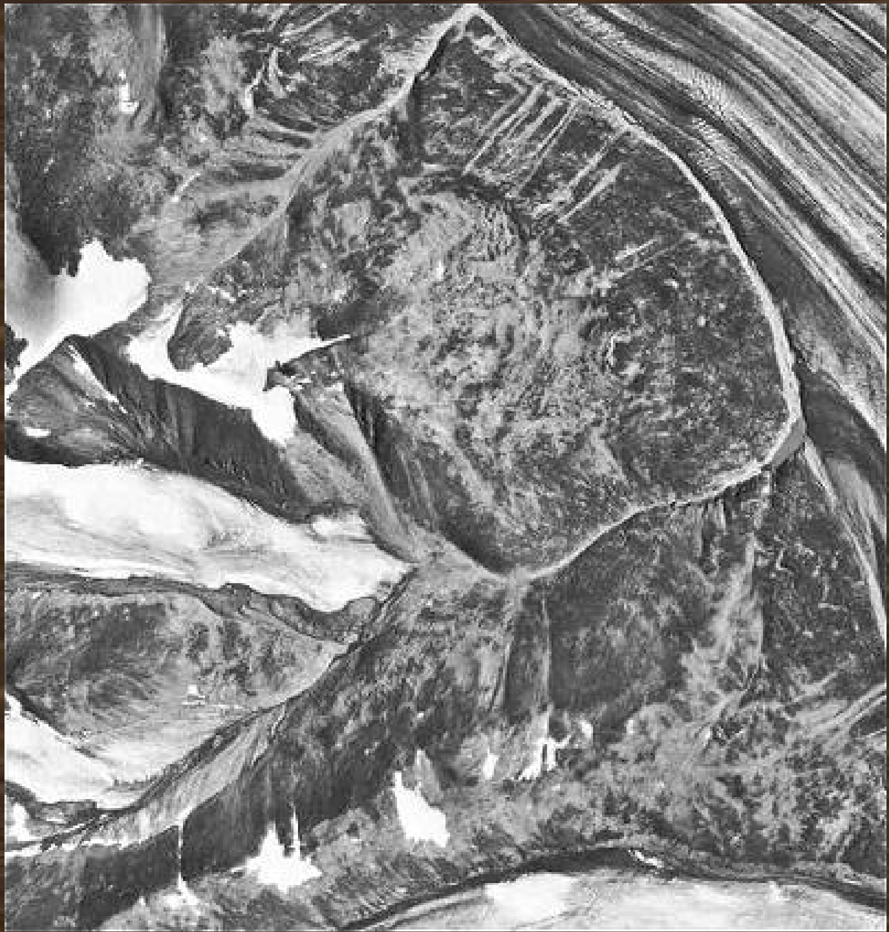
What about the minimum value of depth (object closest to the cameras) ?

$$z_{\min} = B.f / D_{\max}$$

What is D_{max} ? $$D_{\max} = X_{\max}$$

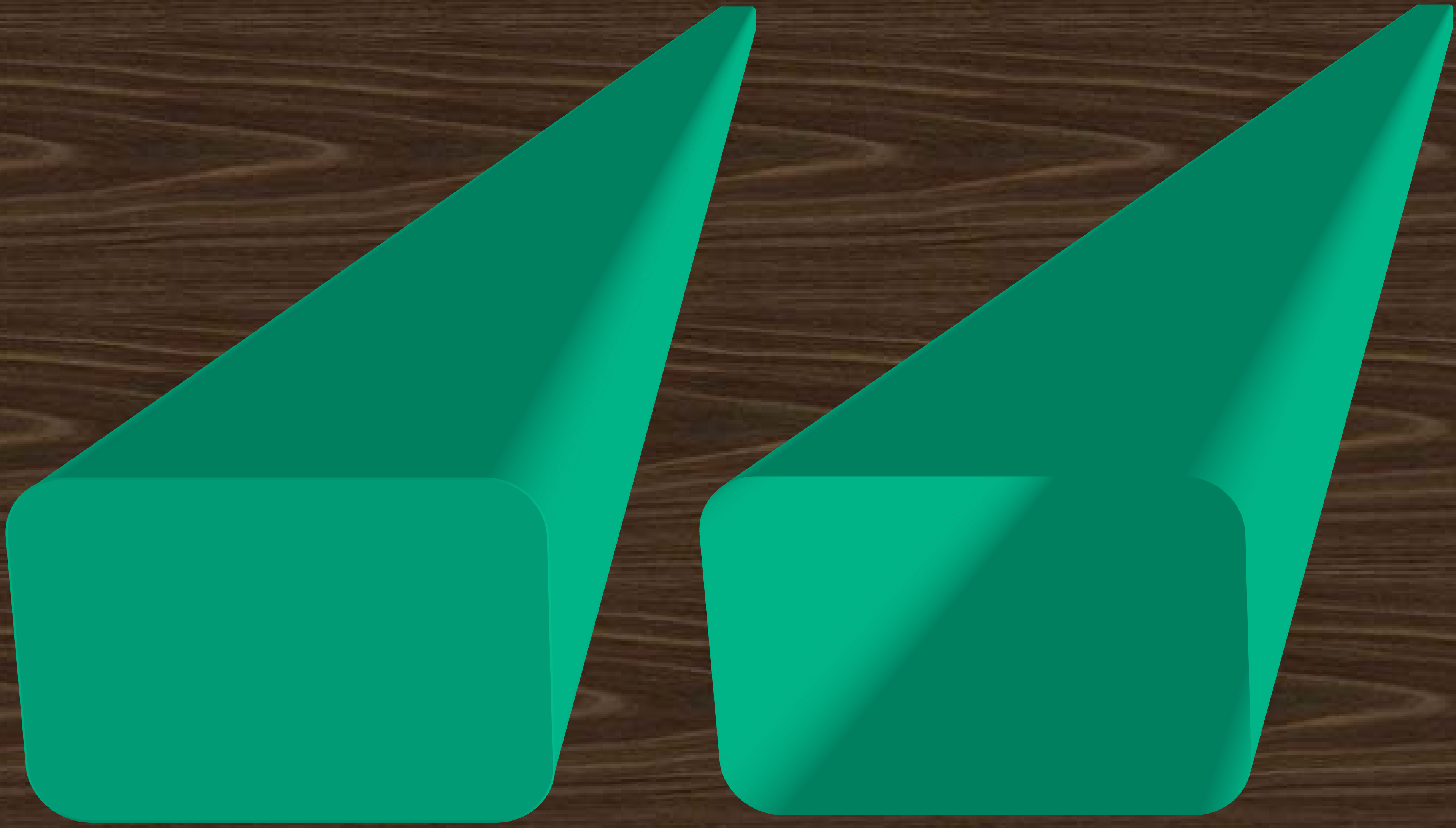
X_{max} depends on f and image resolution (in other words, angle of field-of-view or FOV).



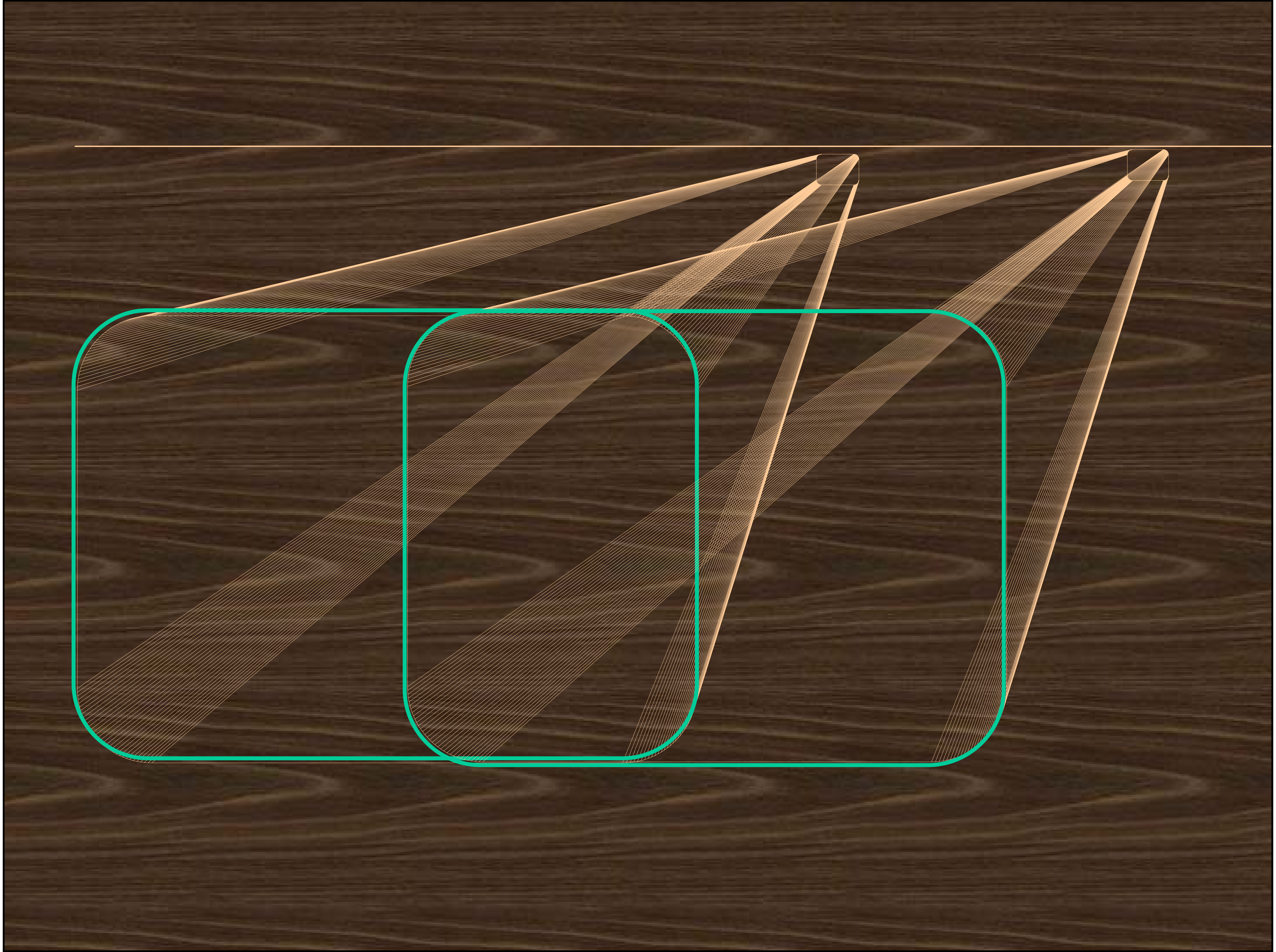
General Stereo Views

Perfect Stereo Views



Perfect Stereo Views

Perfect Stereo Views



We can also have arbitrary pair of views from two cameras.

- The baseline may not lie on any of the principle axis
- The viewing axes of the cameras may not be parallel
- Unequal focal lengths of the cameras
- The coordinate systems of the image planes may not be aligned

Take home exercises/problems:

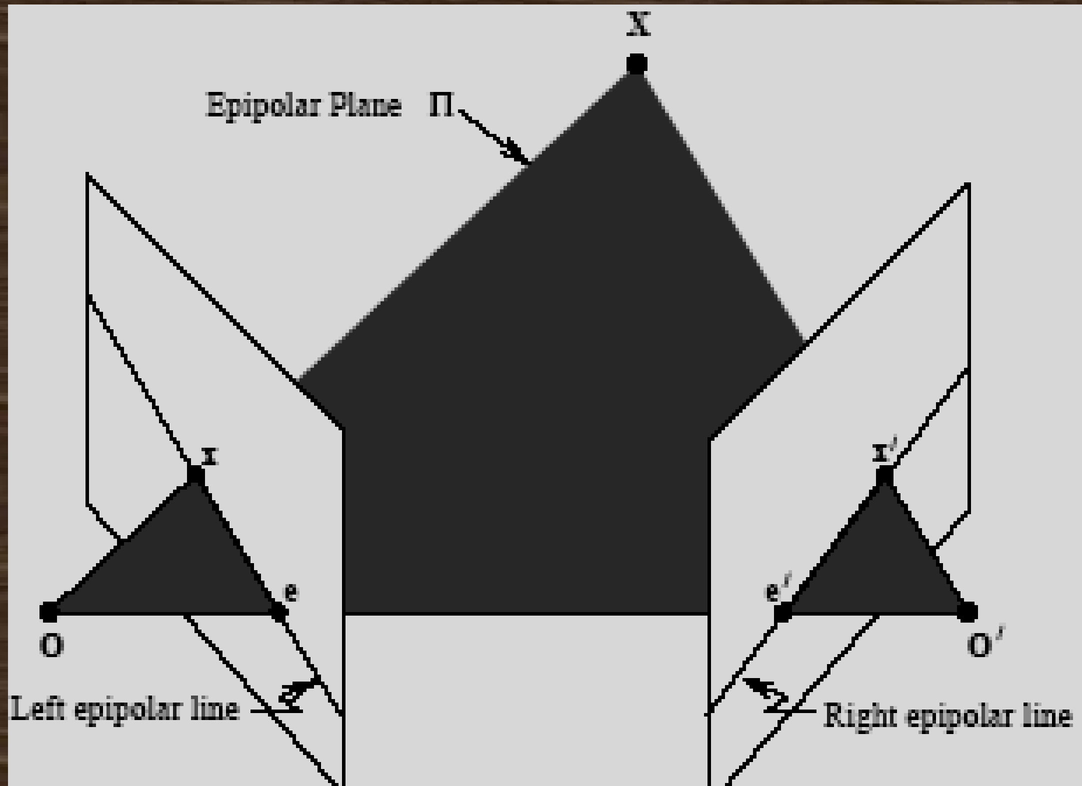
What about Epipolar line in cases above ?

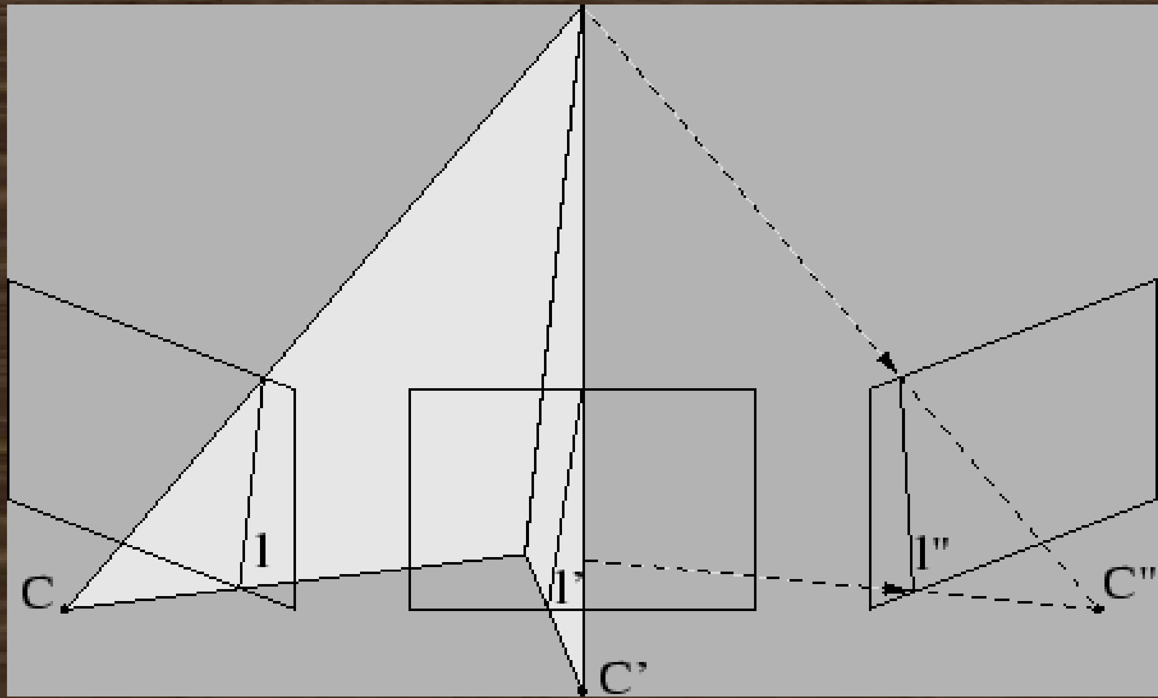
How do you derive the equation of an epipolar line ?

In general we may have multiple views (2 or more) of a scene. Typically used for 3D surveillance tasks.

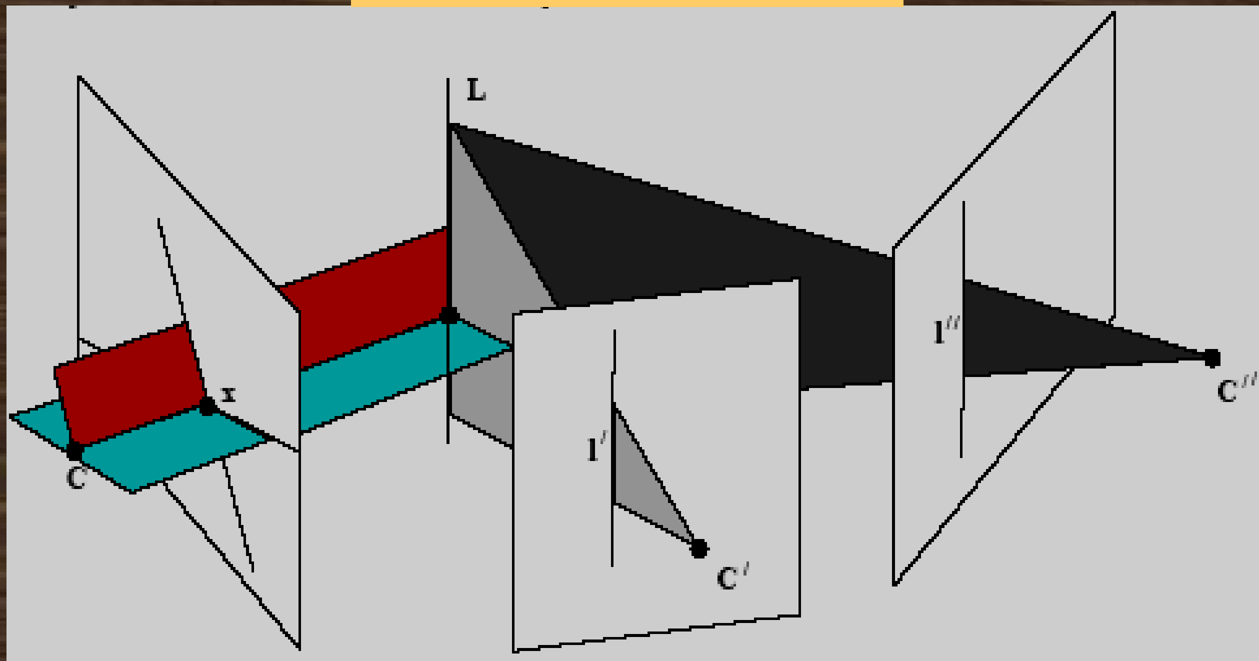
In case of a set of arbitrary views used for 3-D reconstruction (object structure, surface geometry, modeling etc.), methods used involve:

- KLT (Kanade-Lucas-Tomasi)- tracker
- Bundle adjustment
- 8-point DLT algorithm
- Zhang's homography
- *Tri-focal tensors*
- Cheriality and DIAC
- Auto-calibration
- Metric reconstruction
- RANSAC





Tri-focal tensors



End of Lectures on -

Transformations,
Imaging Geometry
and
Stereo Vision