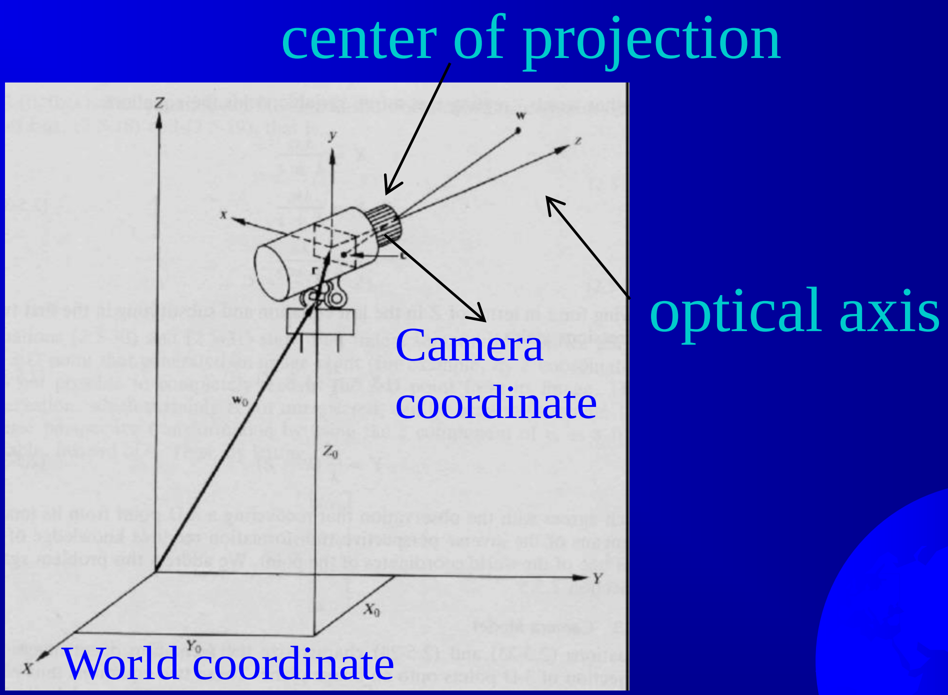
Computer Vision
---*Camera Calibration I*

Dr. WU Xiaojun

2020.9.16

World and Camera coordinate systems

- In general, the world and camera coordinate systems are not aligned.

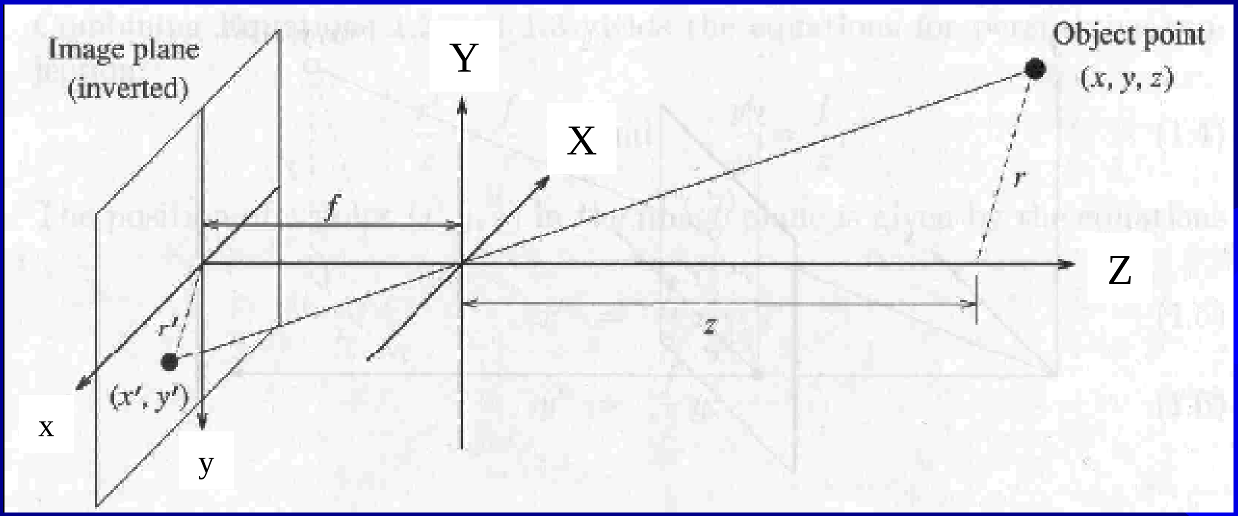


World and Camera coordinate systems (cont'd)

- To simplify mathematics, let's assume:

  (1) The center of projection coincides with the origin of the world coordinate system.

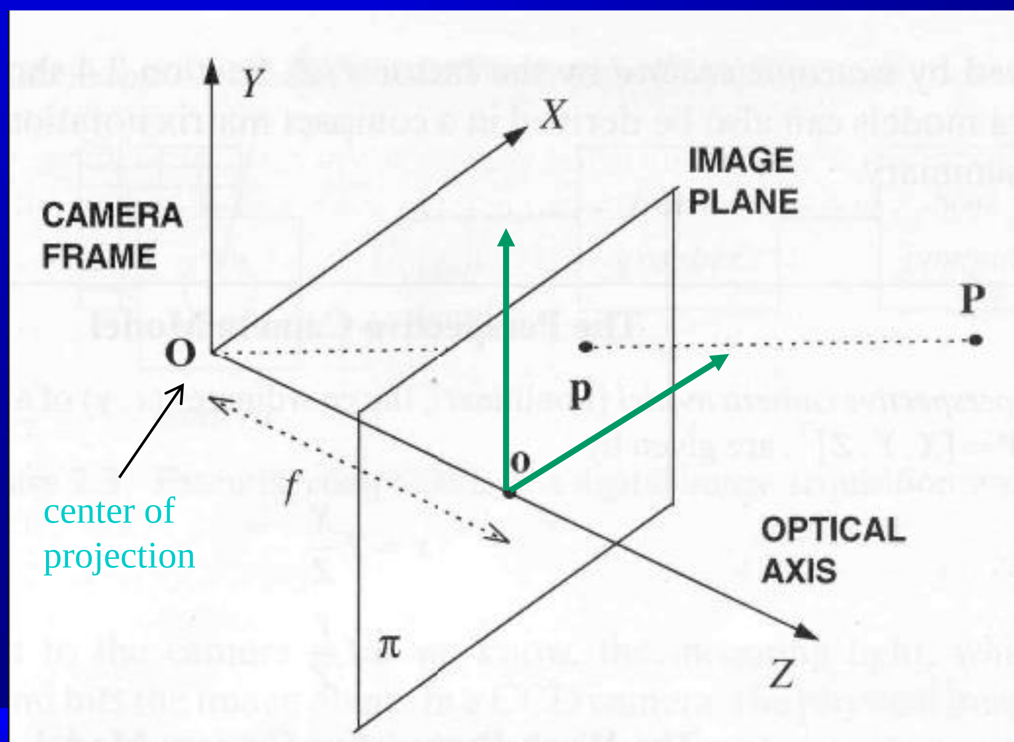
  (2) The optical axis is aligned with the world's z-axis and x,y are parallel with X, Y



World and Camera coordinate systems (cont'd)

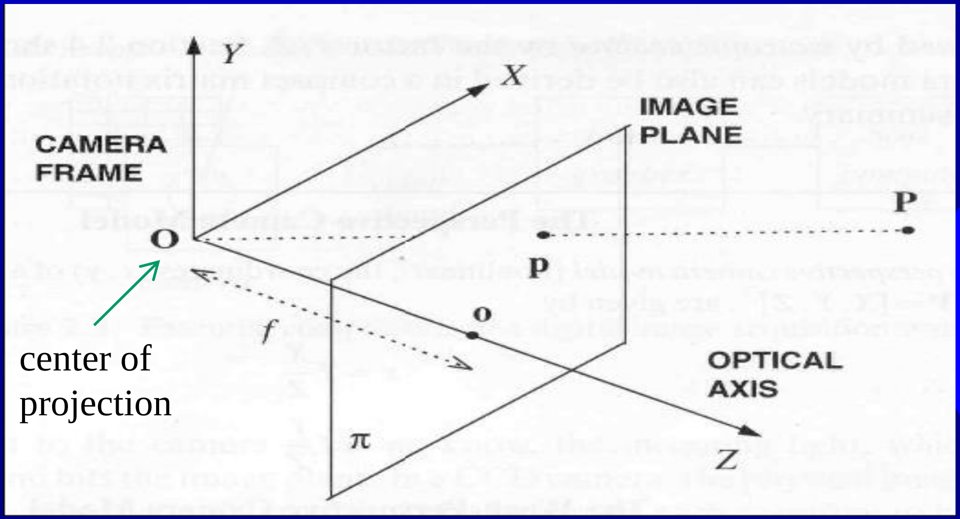
(3) Avoid image inversion by assuming that the image plane is in front of the center of projection.

(4) The origin of the image plane is the principal point.



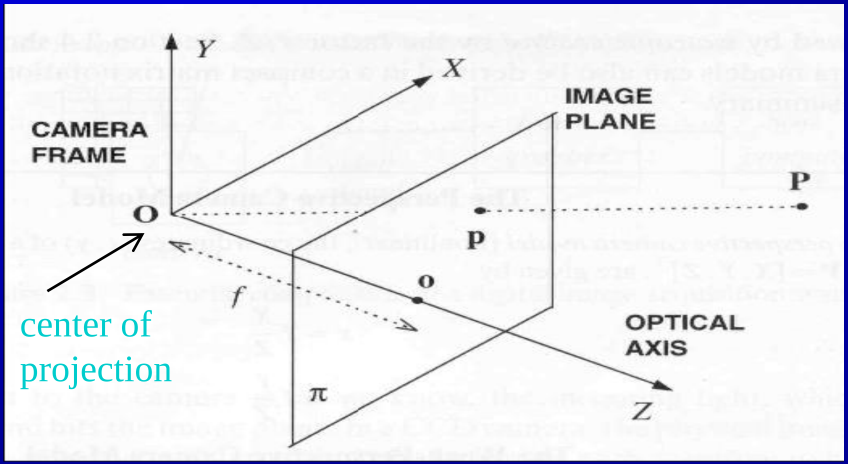
Terminology - Summary

- The model consists of a plane (image plane) and a 3D point *O* (*center of projection*).
- The distance *f* between the image plane and the center of projection *O* is the *focal length* (e.g., the distance between the lens and the CCD array).



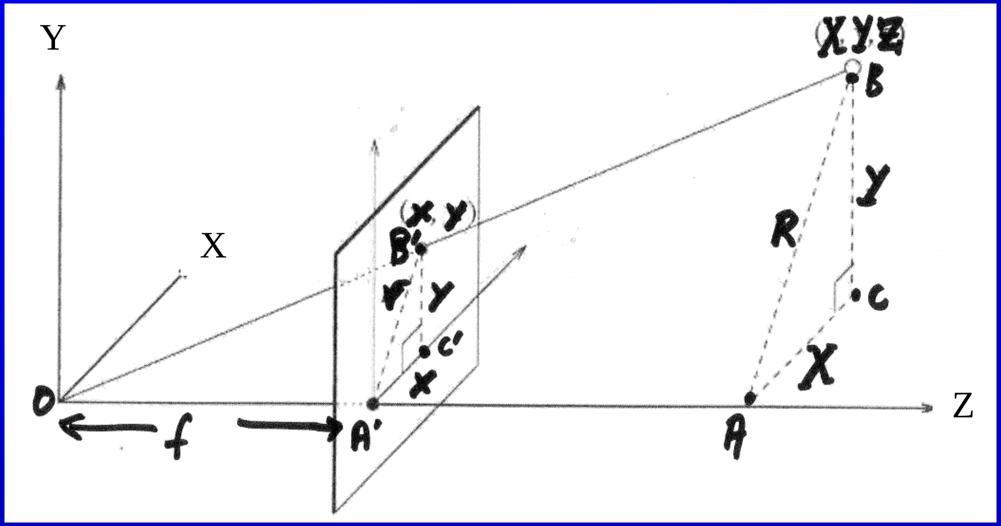
Terminology - Summary (cont'd)

- The line through *O* and perpendicular to the image plane is the *optical axis*.
- The intersection of the optical axis with the image plane is called *principal point*.



<u>Note:</u> the principal point is not necessarily the image center.

The equations of perspective projection



(1) from $OA'B'$ and OAB: $\frac{f}{Z} = \frac{r}{R}$

(2) from $A'B'C'$ and ABC: $\frac{x}{X} = \frac{y}{Y} = \frac{r}{R}$

$$\frac{x}{X} = \frac{y}{Y} = \frac{f}{Z}$$

perspective proj. eqs: $x = \frac{Xf}{Z}$ $\quad y = \frac{Yf}{Z}$ $\quad z = f$

The equations of perspective projection (cont'd)

- Using matrix notation (homogeneous):

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & f & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

or

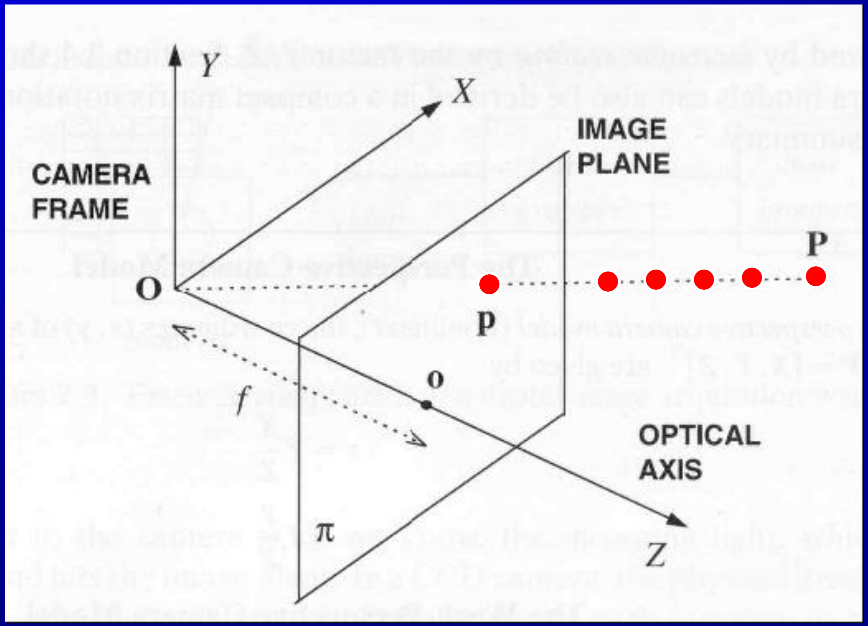
$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Verify the correctness of the above matrix
  - homogenize using $w = Z$

$$x = \frac{x_h}{w} = \frac{fX}{Z} \qquad y = \frac{y_h}{w} = \frac{fY}{Z} \qquad z = \frac{z_h}{w} = f$$

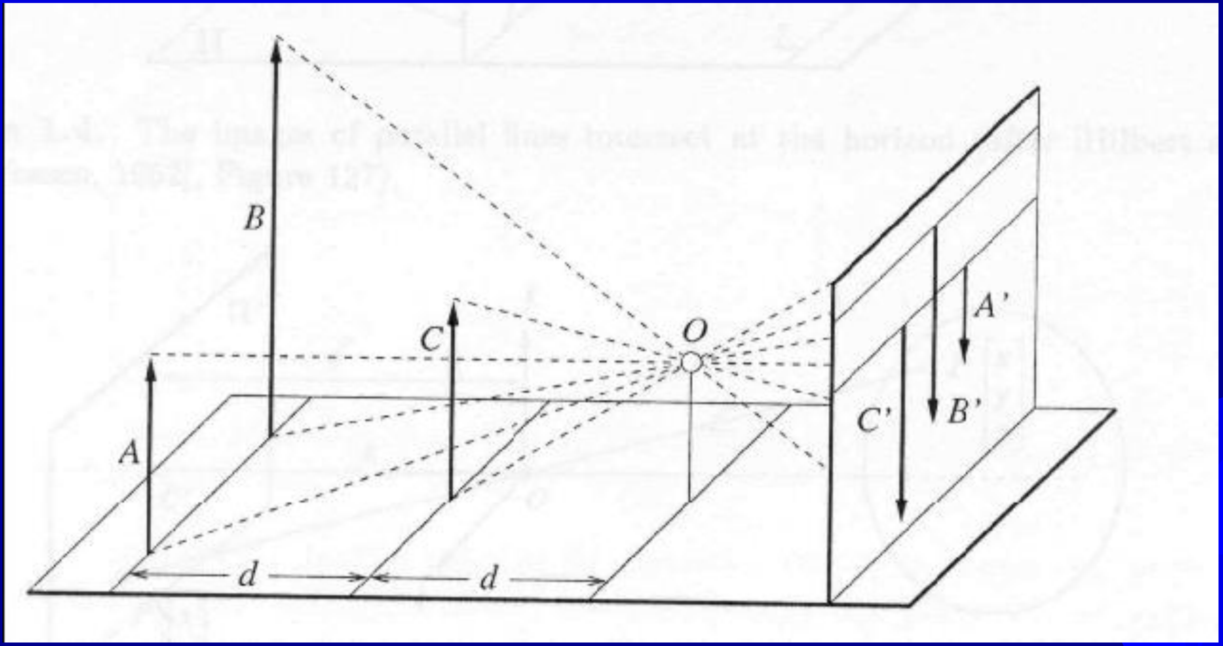
Properties of perspective projection

- Many-to-one mapping
  - The projection of a point is *not* unique
  - Any point on the line *OP* has the same projection



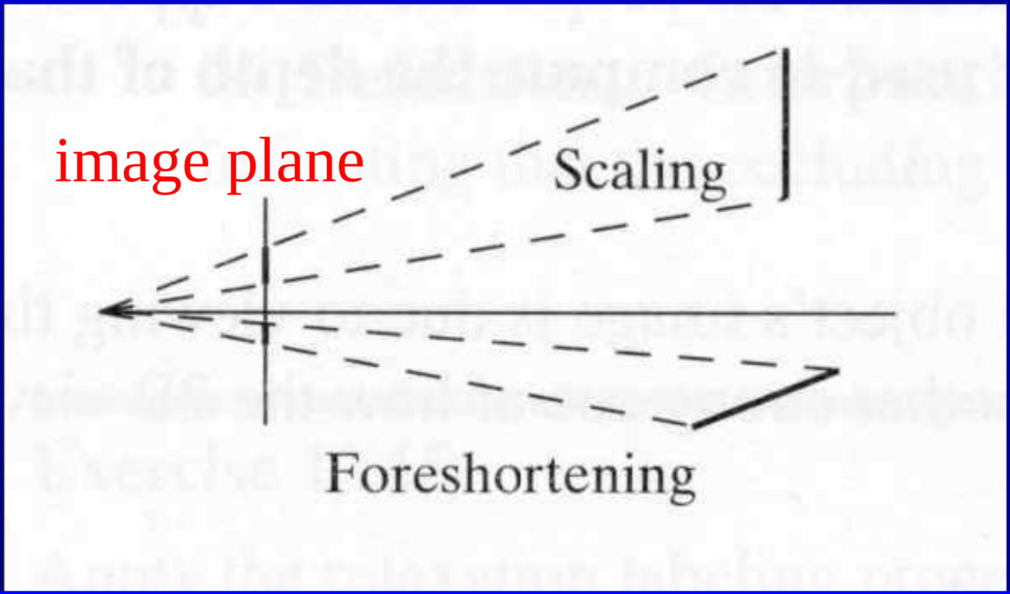
Properties of perspective projection (cont'd)

- Scaling/Foreshortening
  - Object's image size is inversely proportional to the distance of the object from the camera.



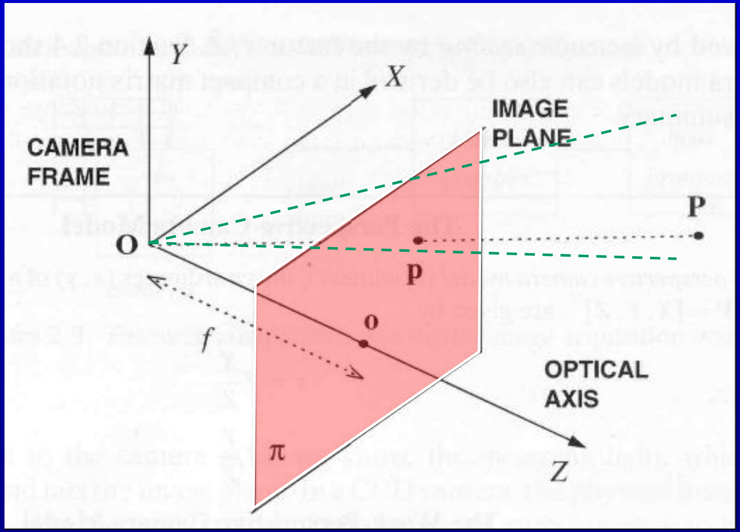
Properties of perspective projection (cont'd)

- When a line (or surface) is parallel to the image plane, the effect of perspective projection is *scaling*.
- When an line (or surface) is not parallel to the image plane, the effect is *foreshortening* (i.e., perspective distortion).



Properties of perspective projection (cont'd)

- Effect of focal length
  - As *f* gets smaller, more points project onto the image plane (*wide-angle camera*).
  - As *f* gets larger, the field of view becomes smaller (more *telescopic*).



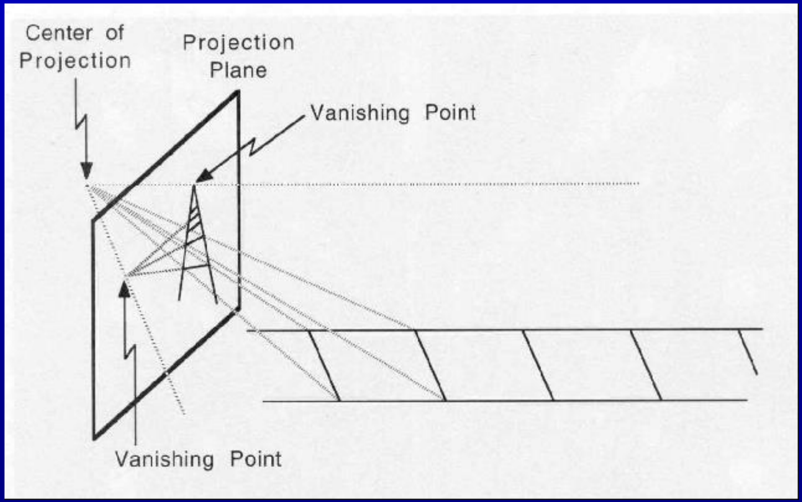
$$x = \frac{x_h}{w} = \frac{fX}{Z} \qquad y = \frac{y_h}{w} = \frac{fY}{Z} \qquad z = \frac{z_h}{w} = f$$

Properties of perspective projection (cont'd)

- What happens to lines, distances, angles and parallelism?
  - Lines in 3D project to lines in 2D (with an exception …)
  - Distances and angles are *not* preserved.
  - Parallel lines *do not* in general project to parallel lines due to foreshortening (unless they are parallel to the image plane).

Properties of perspective projection (cont'd)

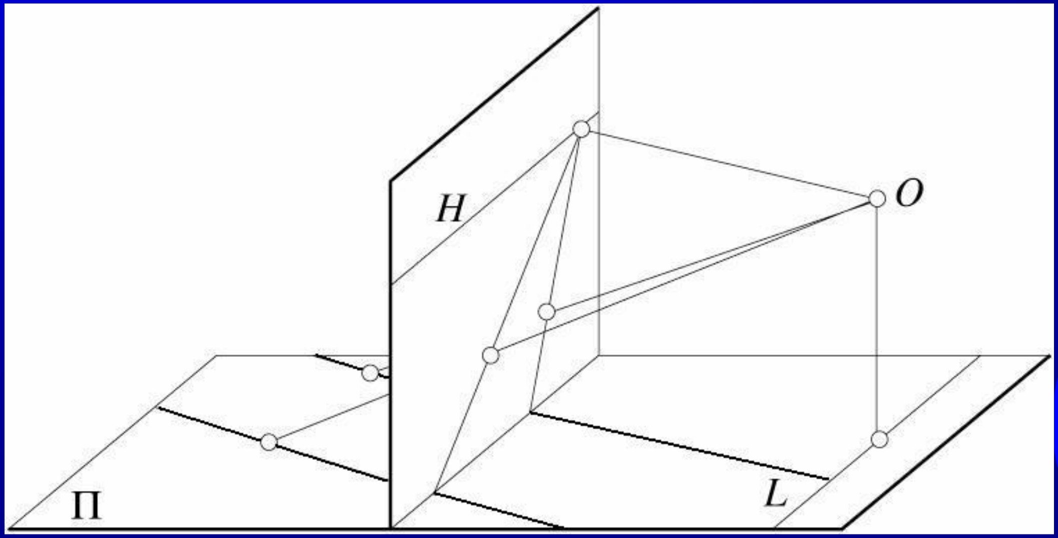
- Vanishing point(消失点):
  - Parallel lines in space project perspectively onto lines that on extension intersect at a single point in the image plane called *vanishing point* (or *point at infinity*).
  - The vanishing point of a line depends on the **orientation** of the line and **not on the position** of the line.



Note: vanishing points might lie outside of the image plane!

Properties of perspective projection (cont'd)

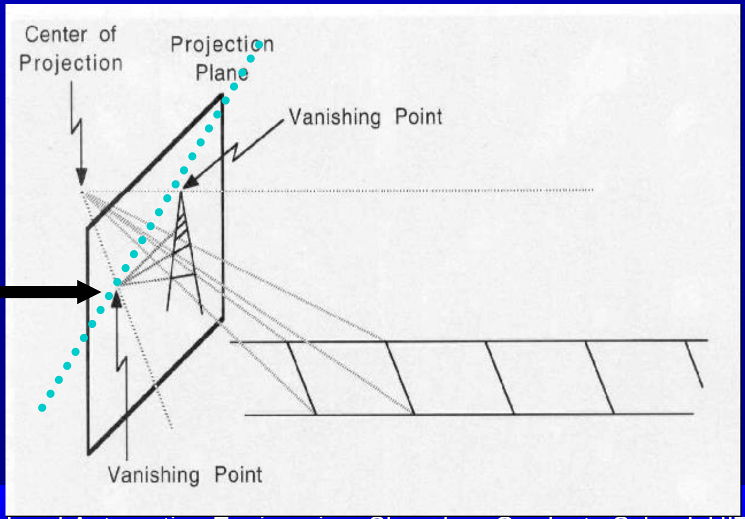
- Alternative definition for vanishing point:
  - The vanishing point of any given line in space is located at the point in the image where a parallel line through the center of projection intersects the image plane. （空间中任何给定直线的灭点都位于图像中通过投影中心的平行线与图像平面相交的点上）



Properties of perspective projection (cont'd)

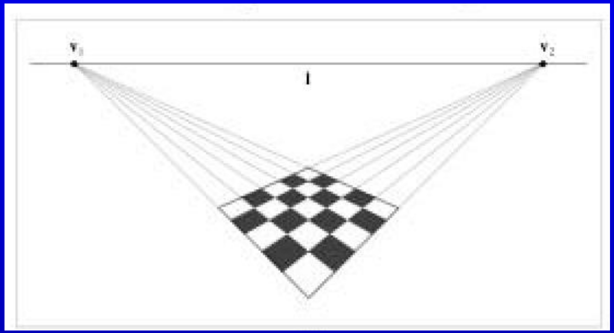
- Vanishing line:
  - The vanishing points of all the lines that lie **on the same plane** form the *vanishing line*.
  - To put it simply, the vanishing line is obtained by the intersection of the image plane with a plane parallel to the ground, passing through the camera center.

vanishing line

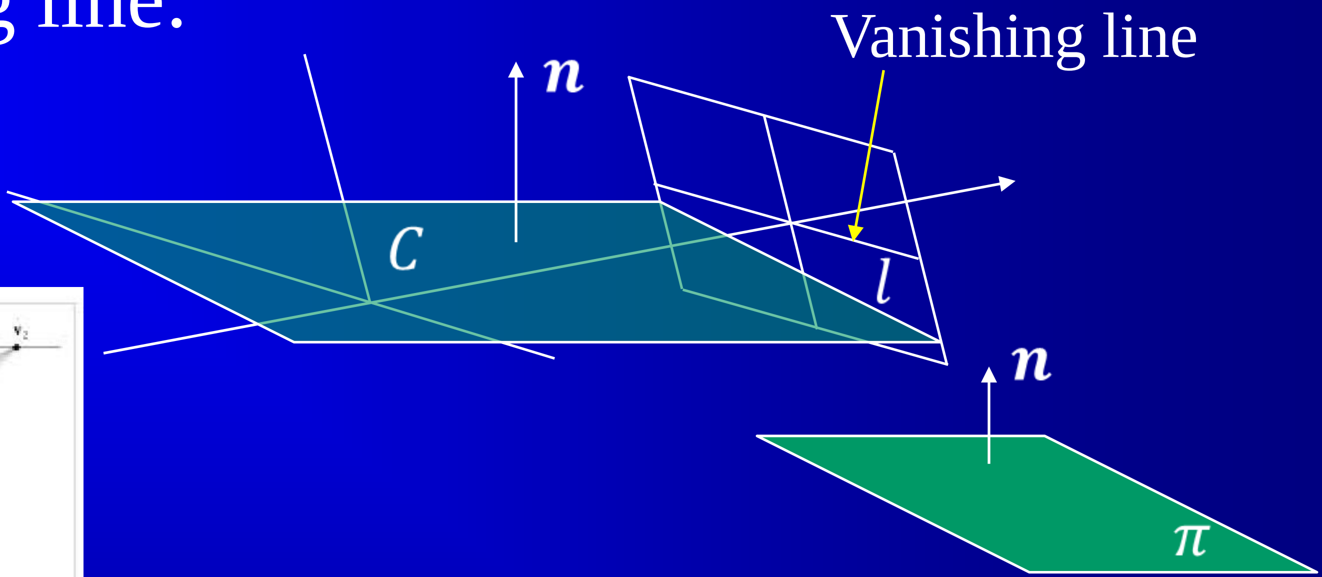


Properties of perspective projection (cont'd)

- Vanishing line:



(a)

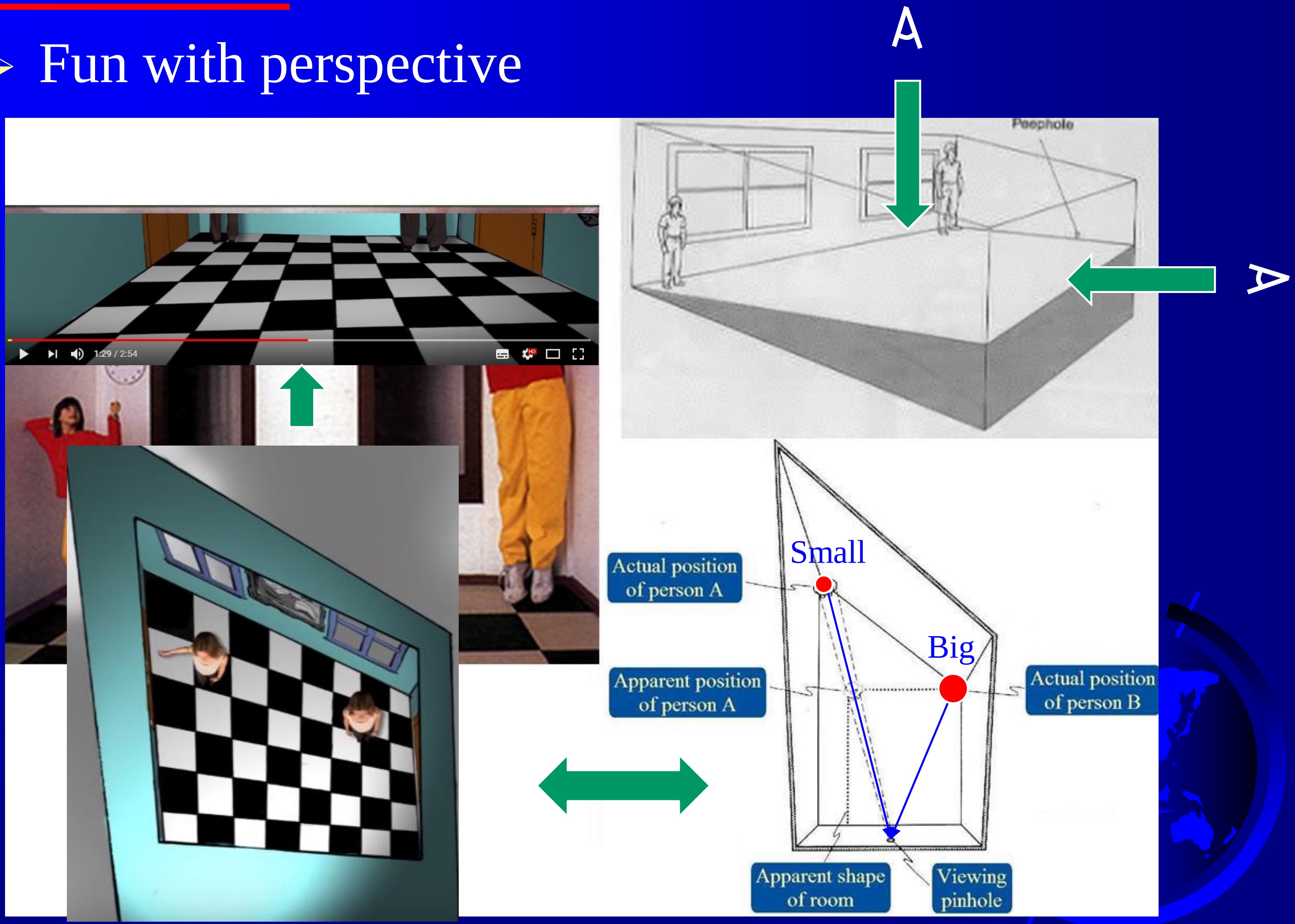


(b)

Vanishing line formation.(a)The two sets of parallel lines on the scene plane comverge tothe vanishing points $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ in the image. The line l through $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ is the vanishing line of theplane.(b) The vanishing line l of a plane $\boldsymbol{\pi}$ is obtained by intersecting the image plane with a plane through the camera centre **C** and parallel to $\boldsymbol{\pi}$.

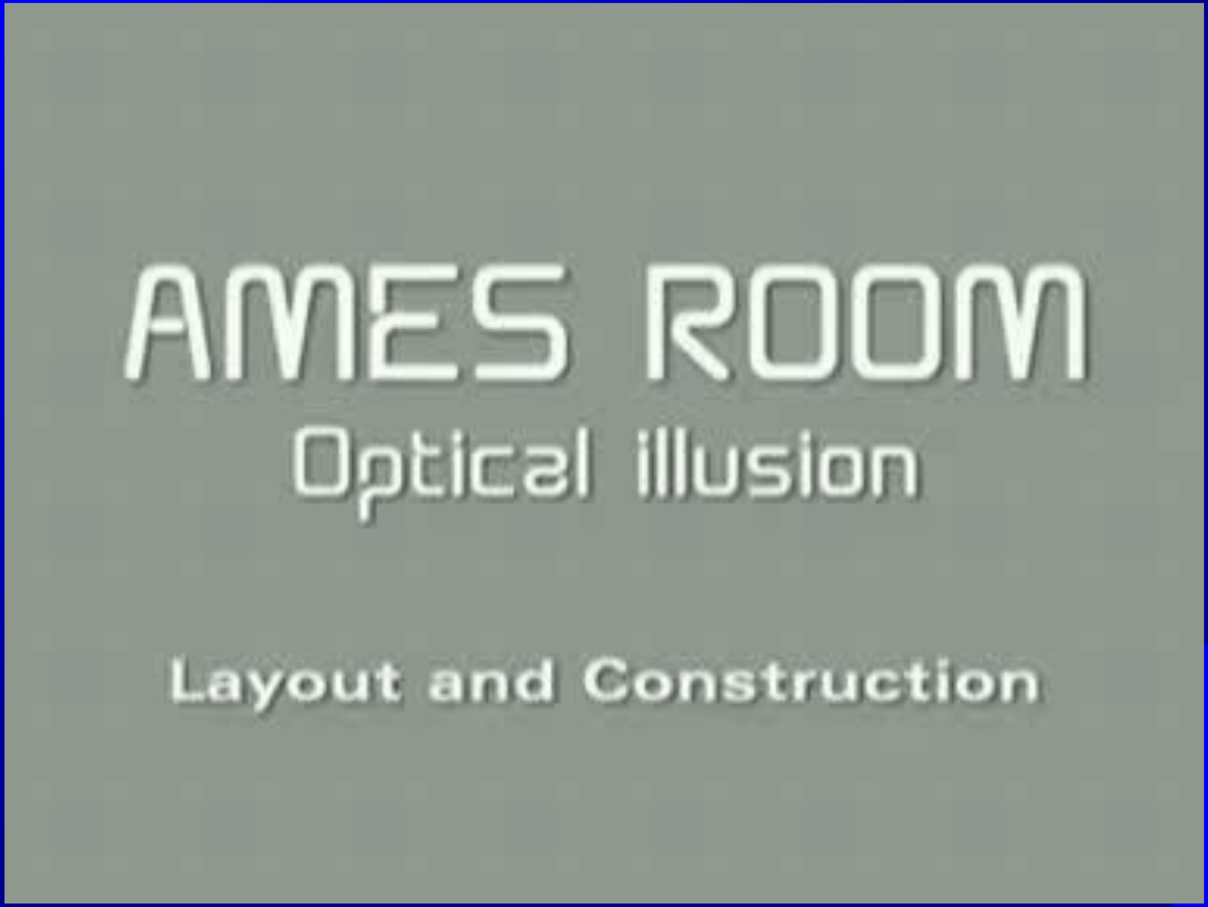
Properties of perspective projection (cont'd)

- Fun with perspective



Properties of perspective projection (cont'd)

- Make a amesh room



Properties of perspective projection (cont'd)

- Fun with perspective

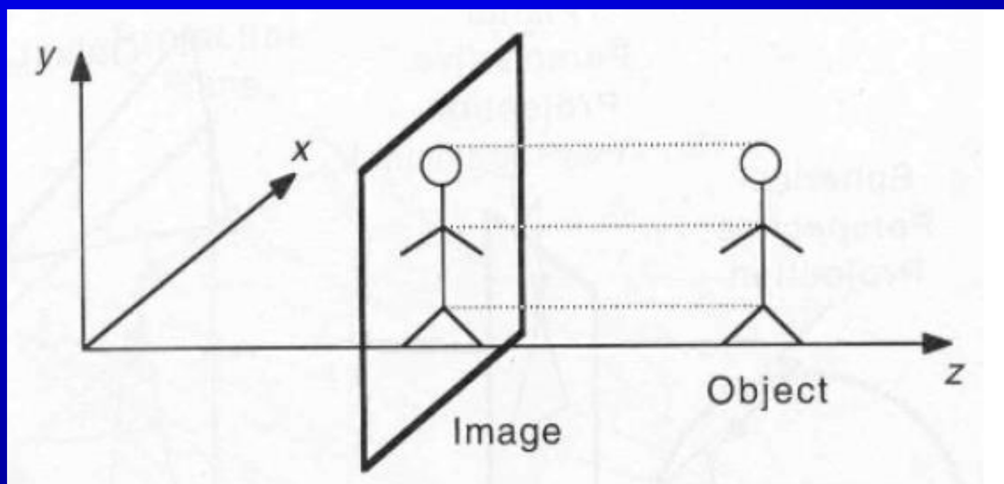
Properties of perspective projection (cont'd)

- Fun with perspective

Orthographic Projection（正交投影）

- The projection of a 3D object onto a plane by a set of parallel rays orthogonal to the image plane.

$$f -> \infty \ (\text{i.e., } f/Z -> 1)$$



$$x = \frac{Xf}{Z} \qquad y = \frac{Yf}{Z} \qquad z = f$$

orthographic proj. eqs: $x = X, \quad y = Y \quad (\text{drop } Z)$

Orthographic Projection (cont'd)

- Using matrix notation:

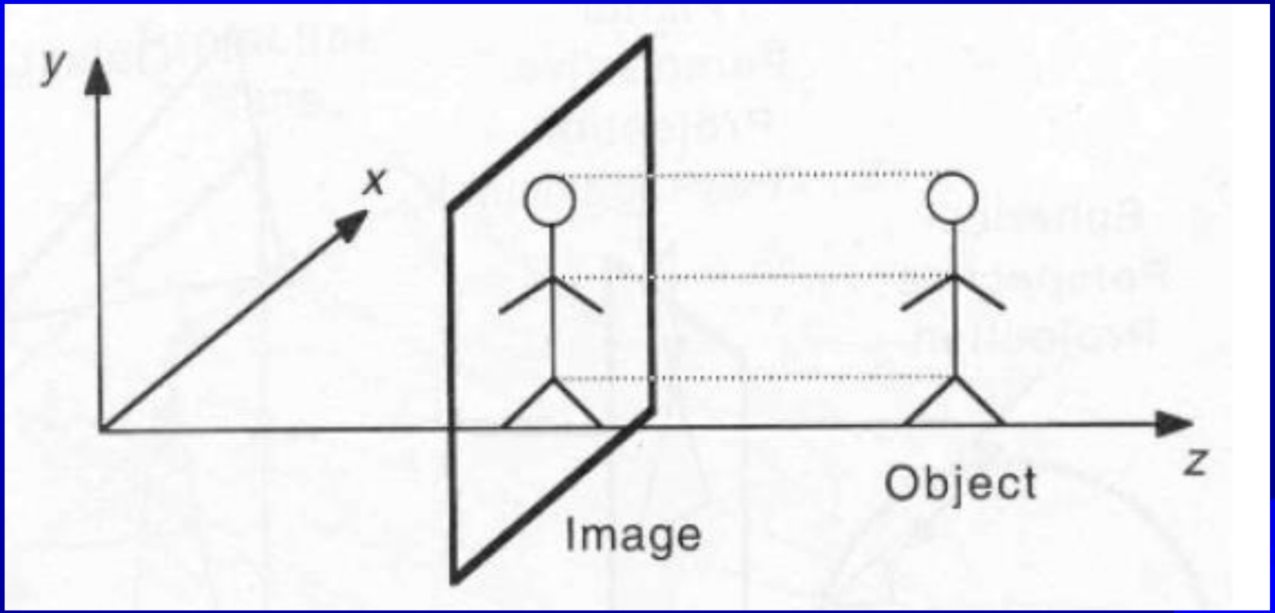
$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

- Verify the correctness of the above matrix (homogenize using w=1):

$$x = \frac{x_h}{w} = X \quad y = \frac{y_h}{w} = Y$$

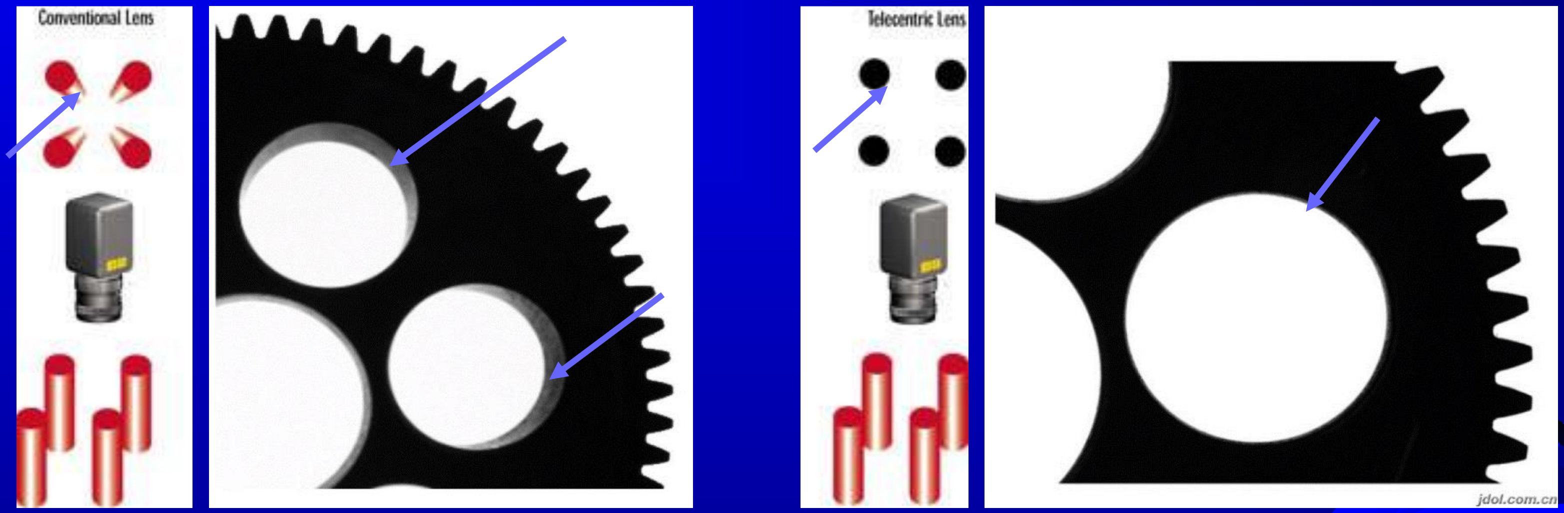
Properties of orthographic projection

- Parallel lines project to parallel lines.
- Size does not change with distance from the camera.



Properties of orthographic projection

- Telecentric Lenses （远心镜头）
  - In many measurement applications, it is highly desirable to have an imaging system that performs a parallel projection because this eliminates the perspective distortions and removes occlusions of objects.



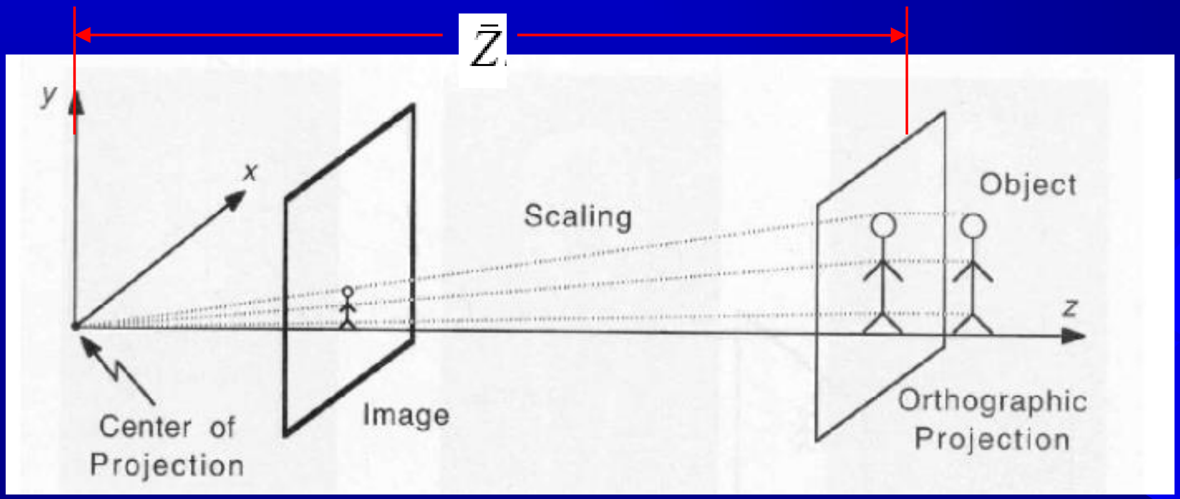
Non-telecentric, Telecentric Imaging results at the image plane of a telecentric and non-telecentric lens system. Notice the telecentric system eliminates perspective distortion.

Weak-perspective projection（弱透视投影）

- Approximate perspective projection by scaled orthographic projection (i.e., linear transformation).
- Good approximation if:

  (1) the object lies close to the optical axis.

  (2) the object's dimensions are small compared to its average distance $\bar{Z}$ from the camera (i.e., $\delta z < \bar{Z}/20$)



Weak perspective projection (cont'd)

weak perspective proj. eqs: $x = \frac{Xf}{Z} \approx \frac{Xf}{\bar{Z}}$ $\quad y = \frac{Yf}{Z} \approx \frac{Yf}{\bar{Z}}$ (drop Z)

- The term $\frac{f}{\bar{Z}}$ is a scale factor now (e.g., every point is scaled by the same factor).

- Using matrix notation:

$$\begin{bmatrix} x_h \\ y_h \\ z_h \\ w \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \bar{Z} \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

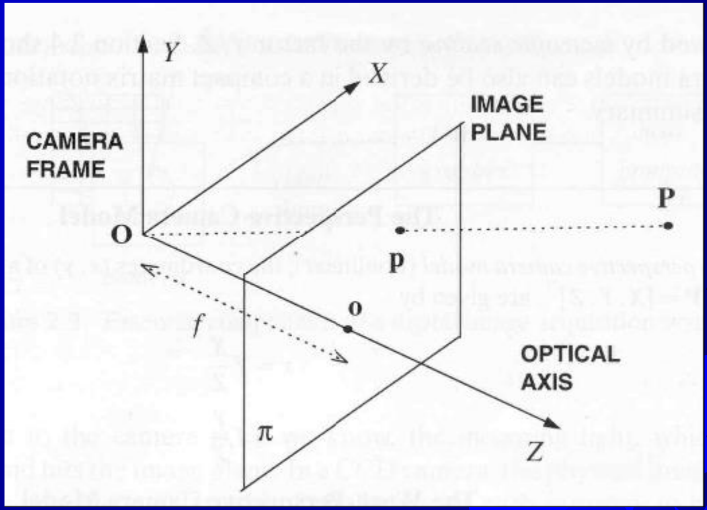
- Verify - homogenize using $w = \bar{Z}$

$$x = \frac{x_h}{w} = \frac{fX}{\bar{Z}} \qquad y = \frac{y_h}{w} = \frac{fY}{\bar{Z}}$$

What assumptions have we made so far?

- Camera and world coordinate systems have been aligned (i.e., all distances are measured in the camera’s reference frame).
- The origin of the image plane is the principal point.

$$x = \frac{x_h}{w} = \frac{fX}{Z} \qquad y = \frac{y_h}{w} = \frac{fY}{Z}$$



World – Pixel Coordinates

- In general, world and pixel coordinates are related by additional parameters such as:

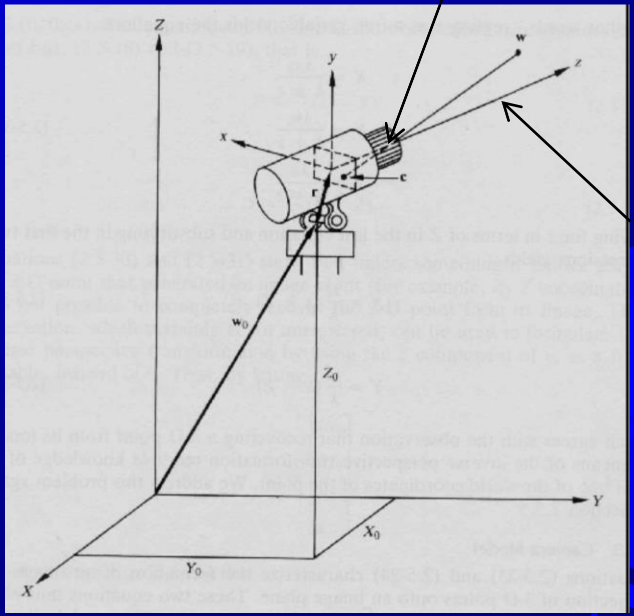
Extri
nsic

  – the position and orientation of the camera

Intri
nsic

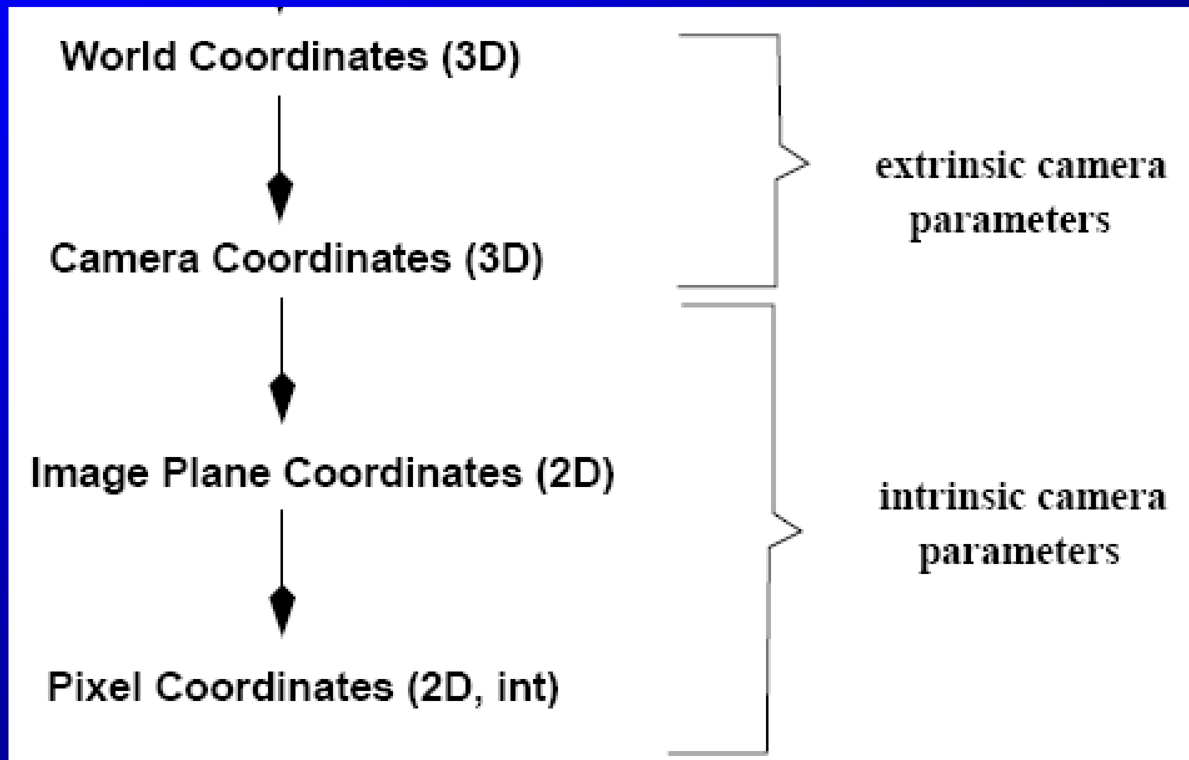
  – the focal length of the lens
  – the position of the principal point
  – the size of the pixels

center of projection



optical axis

Types of parameters

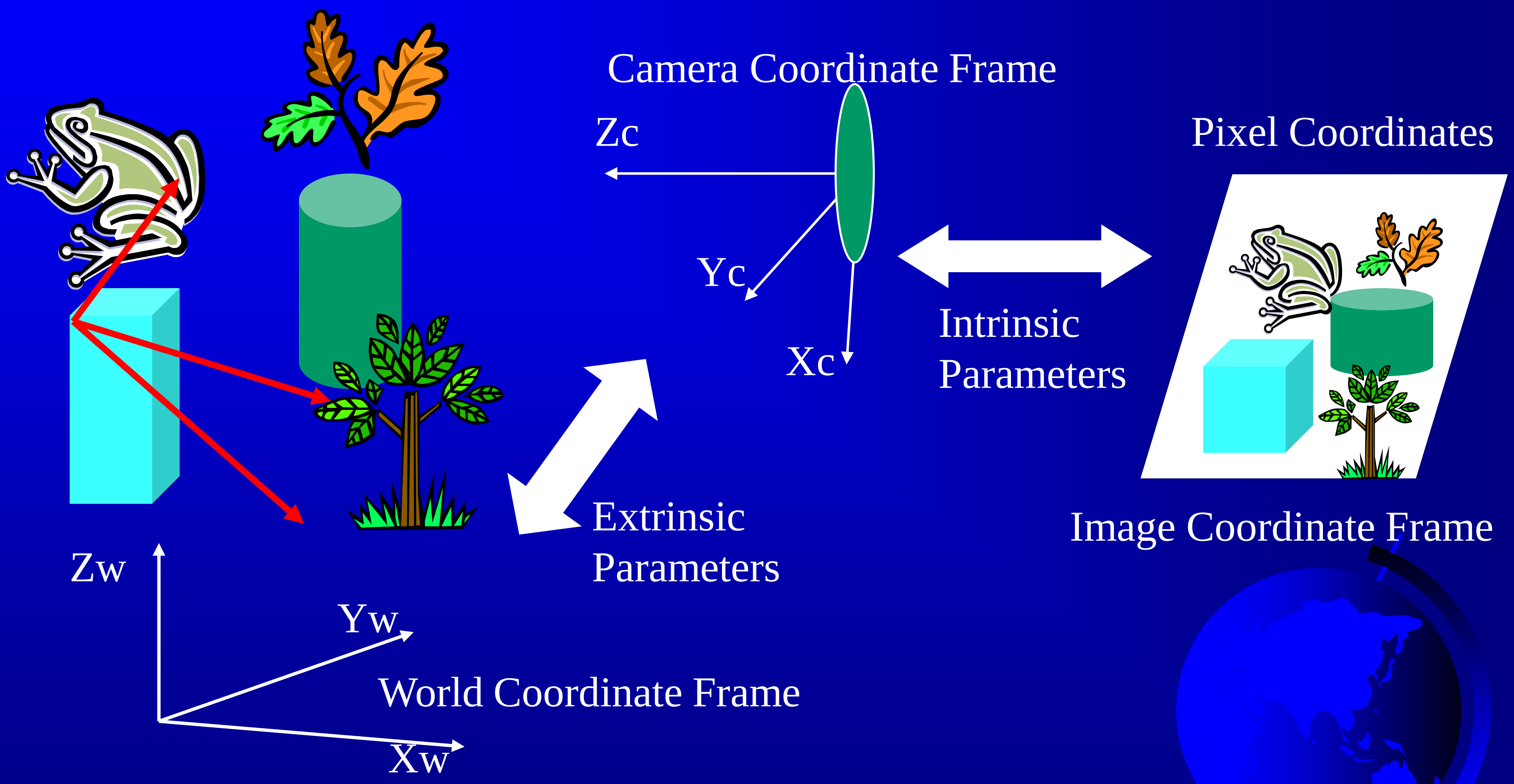


Types of parameters

- Extrinsic: the parameters that define the location and orientation of the camera reference frame with respect to a known world reference frame.

- Intrinsic: the parameters necessary to link the pixel coordinates of an image point with the corresponding coordinates in the camera reference frame.

Types of parameters



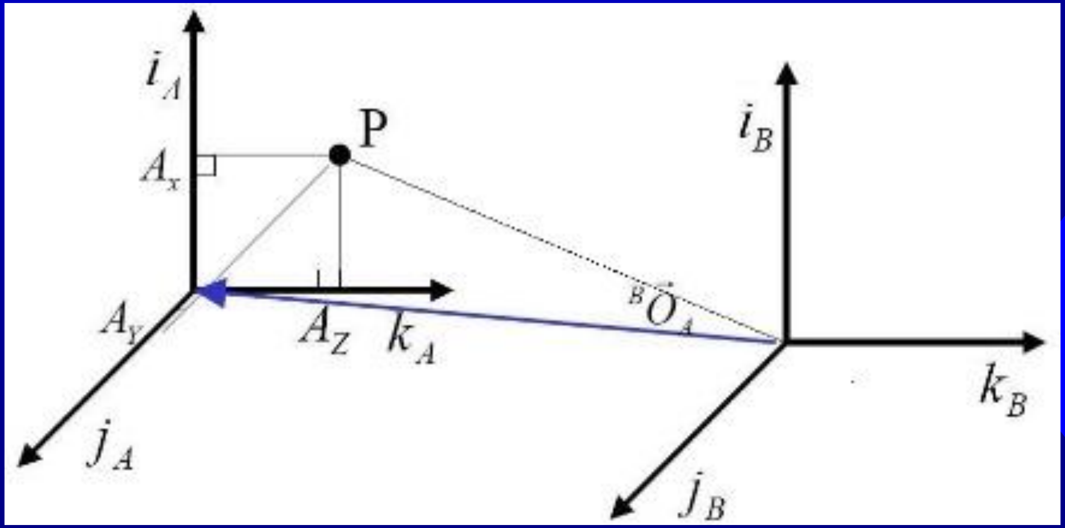
Coordinate Transformation

- Translation

$$^{A}P = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \qquad ^{B}P = \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix}$$

- How does ^{B}P related to ^{A}P?

$$^{B}P = {}^{A}P + {}^{B}O_A.$$



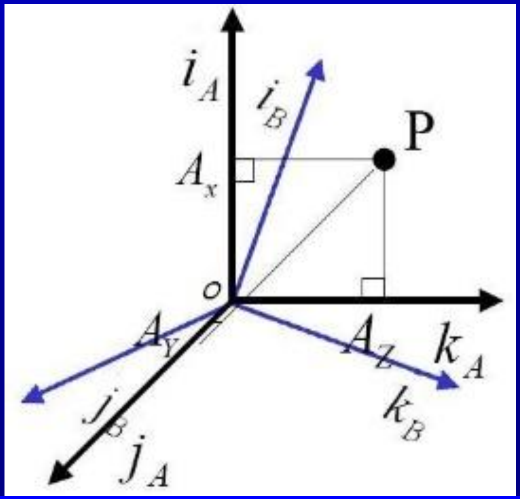
Coordinate Transformation

1. Rotation

$$ {}^{A}P = \begin{pmatrix} A_x \\ A_y \\ A_z \end{pmatrix} \qquad {}^{B}P = \begin{pmatrix} B_x \\ B_y \\ B_z \end{pmatrix} $$

2. How does ${}^{B}P$ related to ${}^{A}P$?

$$ {}^{B}P = {}^{B}_{A}R\,{}^{A}P. $$



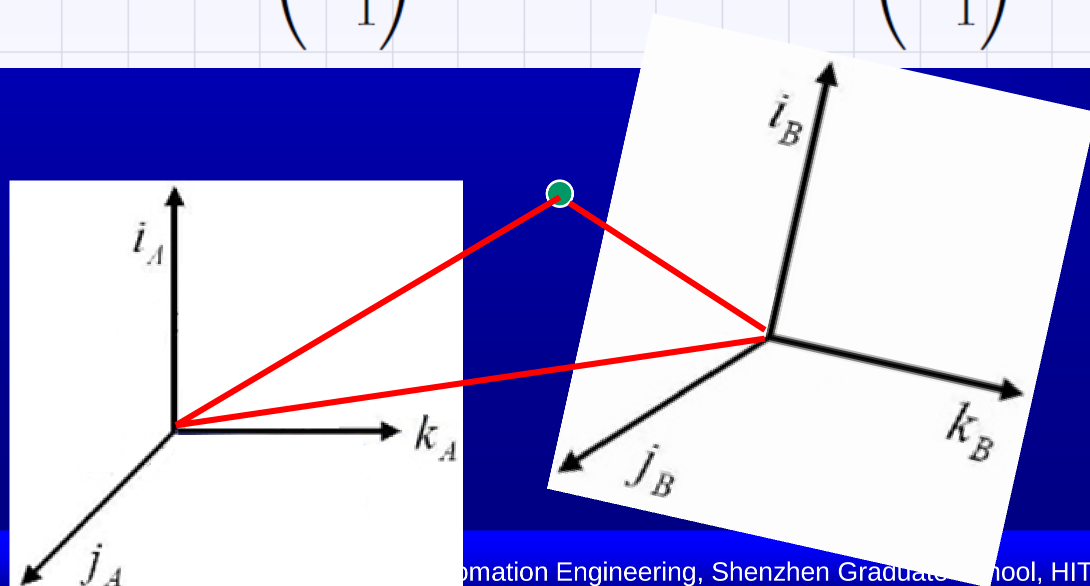
Coordinate Transformation

1. Translation and rotation.
2. Let's write

$$^{B}P = {}^{B}_{A}R\,{}^{A}P + {}^{B}O_{A}$$

as a single matrix equation:

$$\begin{pmatrix} B_x \\ B_y \\ B_z \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{B}_{A}R & {}^{B}\mathbf{O}_{A} \\ \mathbf{0}^{T} & 1 \end{pmatrix} \begin{pmatrix} A_x \\ A_y \\ A_z \\ 1 \end{pmatrix}$$

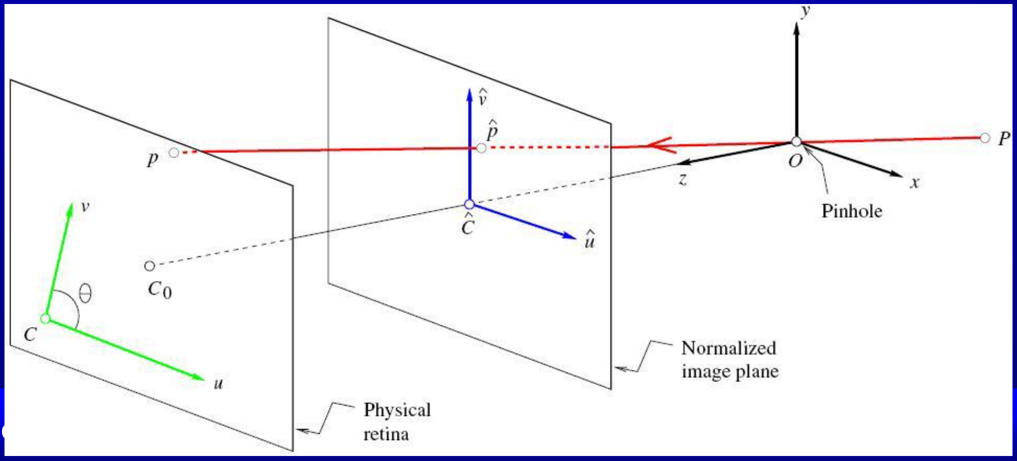


Intrinsic Parameters

- A normalized image plane parallel to its physical retina and locates at a unit distance from the pinhole.
- The perspective projection can be rewritten in the normalized system

$$\begin{cases} \hat{u} = \dfrac{X}{Z} \\ \hat{v} = \dfrac{Y}{Z} \end{cases} \Longleftrightarrow \hat{p} = \frac{1}{Z}(Id \quad 0)\begin{pmatrix} P \\ 1 \end{pmatrix}$$

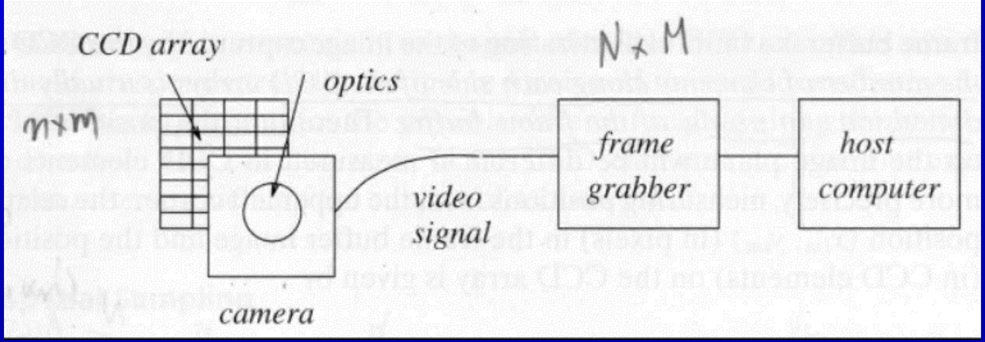
where $\hat{p} \stackrel{\text{def}}{=} (\hat{u}, \hat{v}, 1)^T$ is projection $\hat{p}$ of the point P into the normalized image plane.



Intrinsic Parameters

- On pixel image plane, $f \neq 1$, (u, v) is in pixel units, pixel are rectangular, so the camera has two additional scale parameters.

$$u = \frac{fX}{Z} \qquad v = \frac{fY}{Z}$$



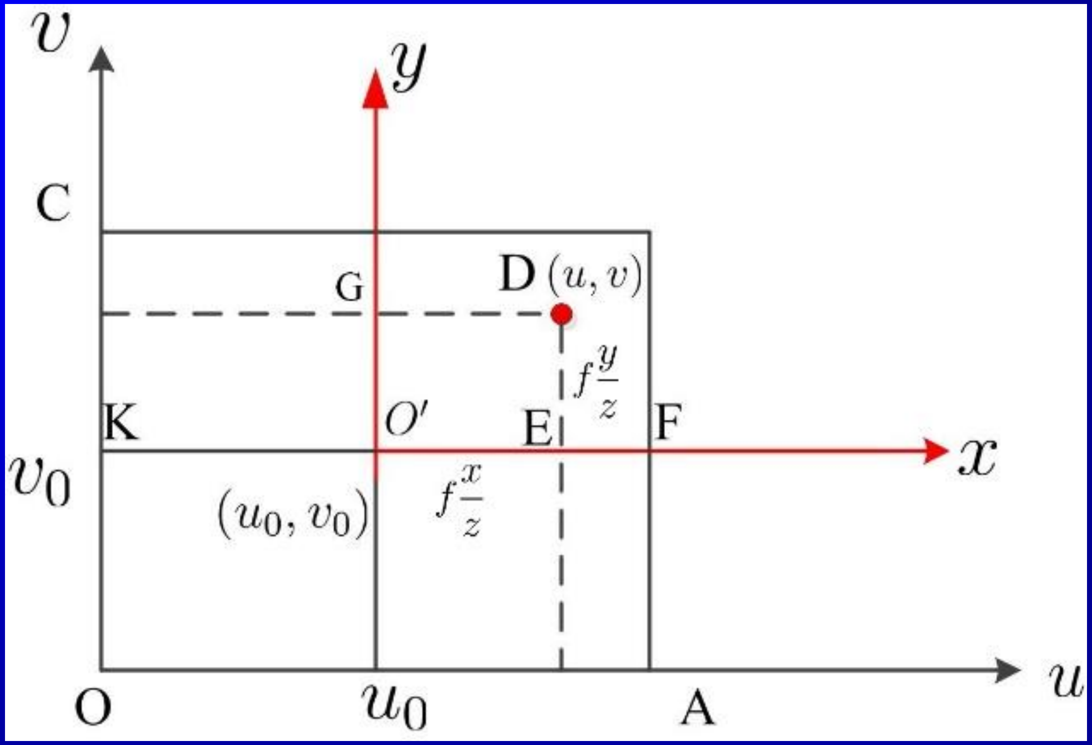
$$\frac{n}{N} \quad \frac{\frac{m}{M}}{} \quad \text{mm/pix}$$

$$u = \frac{fX}{Z} = \frac{f}{n/N}\frac{X}{Z} \qquad v = \frac{fY}{Z} = \frac{f}{m/M}\frac{Y}{Z} \qquad \alpha = \frac{f}{n/N} \qquad \beta = \frac{f}{m/M}$$

$$\boxed{u = \alpha\frac{X}{Z} \qquad v = \beta\frac{Y}{Z}} \Rightarrow \boxed{u = \alpha\hat{u} \qquad v = \beta\hat{v}}$$

Intrinsic Parameters

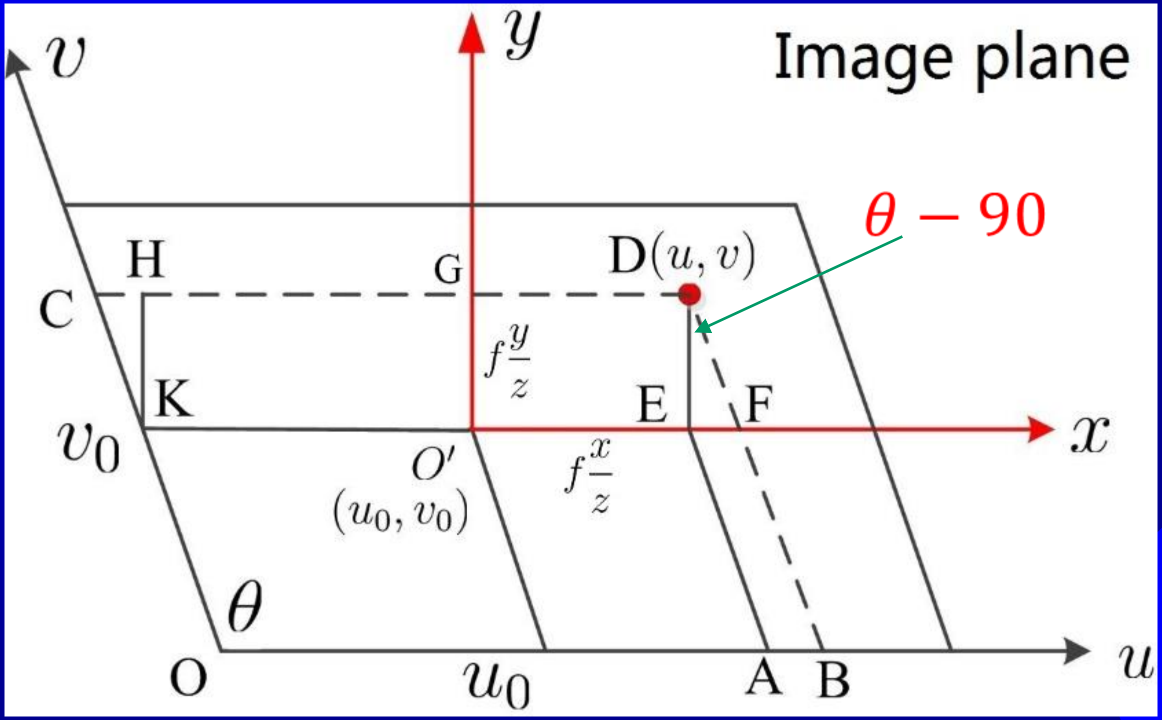
- Transforming the origin of the camera coordinate system.



$$\begin{cases} u = \alpha \dfrac{X}{Z} + u_0 \\ v = \beta \dfrac{Y}{Z} + v_0 \end{cases} \qquad \begin{cases} u = \alpha \hat{u} + u_0 \\ v = \beta \hat{v} + v_0 \end{cases}$$

Intrinsic Parameters

- Camera coordinate may be skewed (倾斜) due to some errors. Angle θ is not equal to 90 degree.



Intrinsic Parameters

❶ Projection

$$\begin{cases} u = \alpha \dfrac{X}{Z} - \alpha \cot\theta \dfrac{Y}{Z} + u_0 \\ v = \dfrac{\beta}{\sin\theta} \dfrac{Y}{Z} + v_0 \end{cases}$$

❷ Change in coordinates between the *physical image frame* and *the normalized one* as a planar affine transformation.

$$\boldsymbol{p} = \mathcal{K}\hat{\boldsymbol{p}}$$

$$\text{where } \boldsymbol{p} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \quad \text{and} \quad \mathcal{K} \stackrel{\text{def}}{=} \begin{pmatrix} \alpha & -\alpha\cot\theta & u_0 \\ 0 & \dfrac{\beta}{\sin\theta} & v_0 \\ 0 & 0 & 1 \end{pmatrix}$$

Intrinsic Parameters

- Put it all together,

$$\boldsymbol{p} = \frac{1}{z}\mathcal{M}\boldsymbol{P}, \quad where \quad \mathcal{M} \stackrel{\text{def}}{=} (\mathcal{K} \quad \mathbf{0})$$

where $\boldsymbol{P}$ denotes denotes the *homogeneous* coordinate vector of P in the camera coordinate system.

Extrinsic Parameters

1. Camera frame (C) is in world frame (W)

$$\begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}R & {}^{C}O_W \\ \mathbf{0}^T & 1 \end{pmatrix} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix}$$

2. Substituting above equation into (20), and yields

$$\boldsymbol{p} = \frac{1}{z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} = \mathcal{K}(\mathcal{R} \quad \boldsymbol{t}) \tag{1}$$

where $\mathcal{R} = {}^{C}_{W}\mathcal{R}$ is a rotation matrix, $\boldsymbol{t} = {}^{C}O_W$ translation vector, and $\boldsymbol{P} = ({}^{W}x, {}^{W}y, {}^{W}z, 1)^T$ HC of P in frame (W).

3. If $\boldsymbol{m}_1^T, \boldsymbol{m}_2^T$ and $\boldsymbol{m}_3^T$ denotes the three rows of $\mathcal{M}$, from (21) $z = \boldsymbol{m}_3 \cdot \boldsymbol{P}$. We rewrite (1) as

$$\begin{cases} u = \dfrac{\boldsymbol{m}_1 \cdot \boldsymbol{P}}{\boldsymbol{m}_3 \cdot \boldsymbol{P}} \\ v = \dfrac{\boldsymbol{m}_2 \cdot \boldsymbol{P}}{\boldsymbol{m}_3 \cdot \boldsymbol{P}} \end{cases}$$

Intrinsic Parameters

- A projection matrix can be written explicitly as a function of its five intrinsic parameters $(\alpha, \beta, u_0, v_0, \theta)$ and its six extrinsic ones (three angles defining $\boldsymbol{R}$) and three components of translation vector $\boldsymbol{t}$.

$$\mathcal{M} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T - \alpha \cot\theta \boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \frac{\beta}{\sin\theta} \boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & \frac{\beta}{\sin\theta} t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{pmatrix}$$

Intrinsic Parameters

- For simplicity, we set $\theta = 90^o$, we get the projection matrix as

$$\mathcal{M} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T + u_0 \boldsymbol{r}_3^T & \alpha t_x + u_0 t_z \\ \beta \boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & \beta t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{pmatrix}$$

where $\boldsymbol{r}_1^T, \boldsymbol{r}_2^T$ and $\boldsymbol{r}_3^T$ are three rows of matrix $\mathcal{R}$. t_x, t_y, t_z are three components of vector $\boldsymbol{t}$.

See You