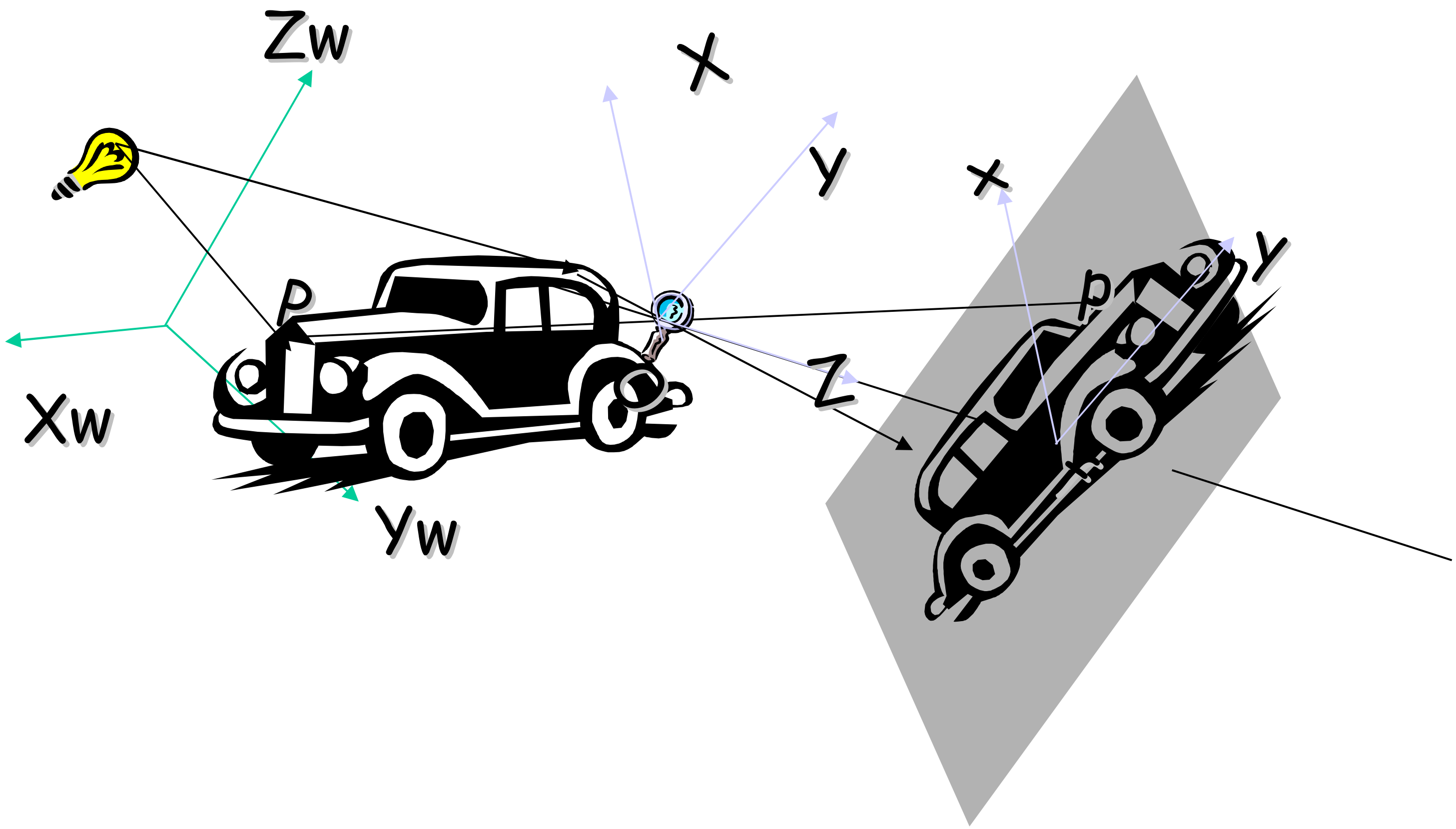
Camera Geometry & Calibration

CS 554 – Computer Vision
Pinar Duygulu
Bilkent University

Coordinate systems



We will use WORLD, CAMERA and Image Coordinate Systems.

Adapted from Octavia Camps

Geometric Camera Models

Issue

- camera may not be at the origin, looking down the z-axis

  extrinsic parameters
- one unit in camera coordinates may not be the same as one unit in world coordinates

  intrinsic parameters

Intrinsic parameters

- Do not depend on the camera location
  - Focal length, CCD dimensions, lens distortion

Extrinsic parameters

- Depend on the camera location
  - Translation, and Rotation parameters

Notions of Geometry

- Homogeneous coordinates
- Matrix representation of geometric transformations
- Extrinsic and intrinsic parameters that relate the world and the camera coordinate frames

Reminder

Dot product

$$\boldsymbol{u} = (u_1, \ldots, u_n)^T$$

$$\boldsymbol{v} = (v_1, \ldots, v_n)$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + \ldots + u_n v_n,$$

$$\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{u}^T \boldsymbol{v} = \boldsymbol{v}^T \boldsymbol{u}$$

When u has unit norm u.v is sign length of projection of v onto u

Cross product

$$\boldsymbol{u} = (u_1, u_2, u_3)^T$$

$$\boldsymbol{v} = (v_1, v_2, v_3)^T$$

$$\boldsymbol{u} \times \boldsymbol{v} \stackrel{\text{def}}{=} \begin{pmatrix} u_2 v_3 - u_3 v_2 \\ u_3 v_1 - u_1 v_3 \\ u_1 v_2 - u_2 v_1 \end{pmatrix}$$

u x v is orthogonal to these two

If u and v have same direction u x v = 0

$$(\boldsymbol{u} \cdot \boldsymbol{v})^2 = |\boldsymbol{u}|^2 |\boldsymbol{v}|^2 \cos^2 \theta,$$

$$|\boldsymbol{v} \times \boldsymbol{v}|^2 = |\boldsymbol{u}|^2 |\boldsymbol{v}|^2 \sin^2 \theta$$

Forsyth & Ponce

Homogeneous coordinates

- Add an extra coordinate and use an equivalence relation
- for 3D
  - equivalence relation k*(X,Y,Z,T) is the same as (X,Y,Z,T)

- Motivation
  - Possible to write the action of a perspective camera as a matrix

Adapted from David Forsyth, UC Berkeley

Homogeneous coordinates

Homogenous/non-homogenous transformations for a 3-d point

- From non-homogenous to homogenous coordinates: add 1 as the 4th coordinate, ie

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

- From homogenous to non-homogenous coordinates: divide 1st 3 coordinates by the 4th, ie

$$\begin{pmatrix} x \\ y \\ z \\ T \end{pmatrix} \rightarrow \frac{1}{T} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

Adapted from Trevor Darrell, MIT

Homogeneous coordinates

Homogenous/non-homogenous transformations for a 2-d point

- From non-homogenous to homogenous coordinates: add 1 as the 3rd coordinate, ie

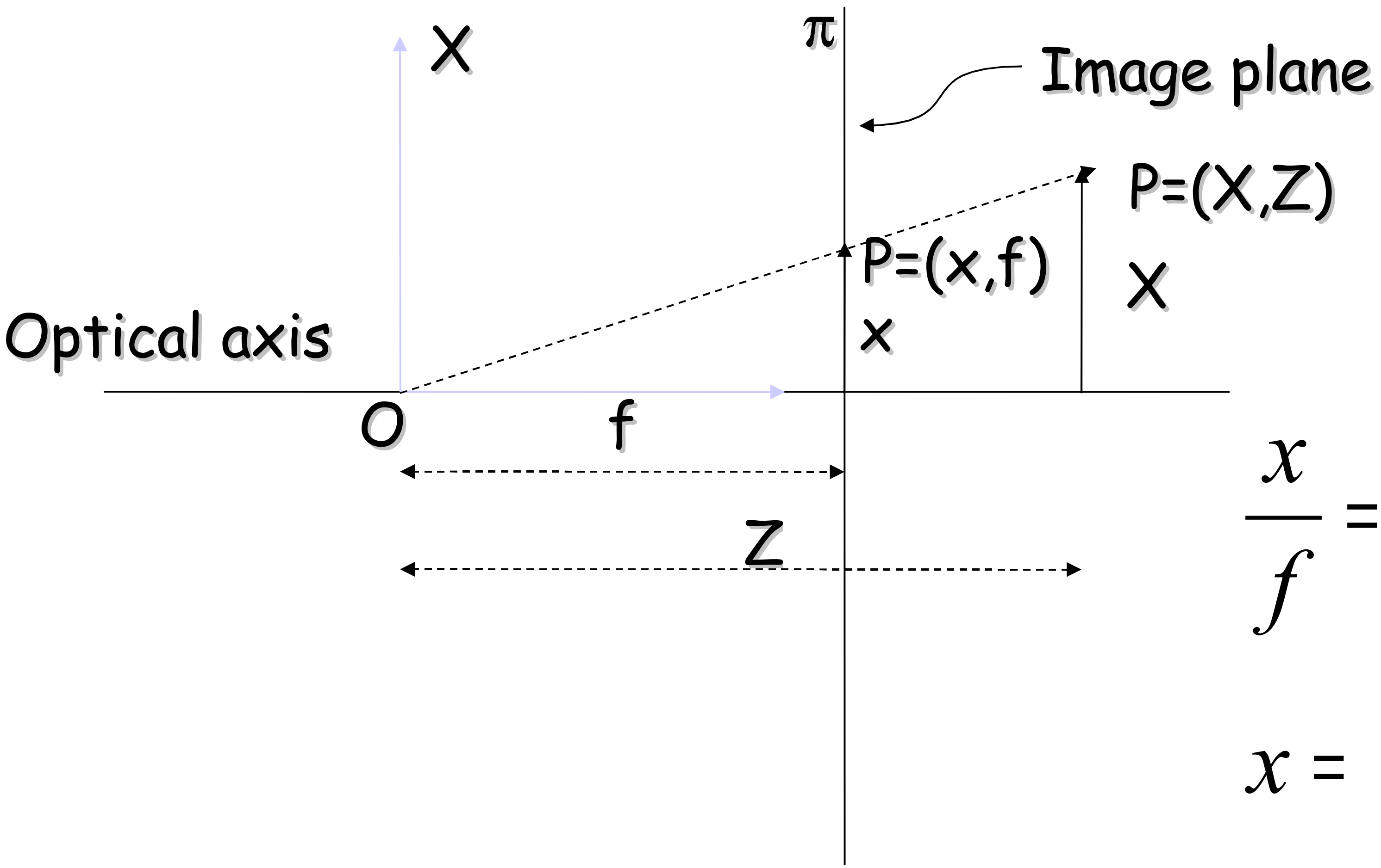
$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$$

- From homogenous to non-homogenous coordinates: divide 1st 2 coordinates by the 3rd, ie

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \rightarrow \frac{1}{z} \begin{pmatrix} x \\ y \end{pmatrix}$$

Adapted from Trevor Darrell, MIT

Pinhole Camera Model

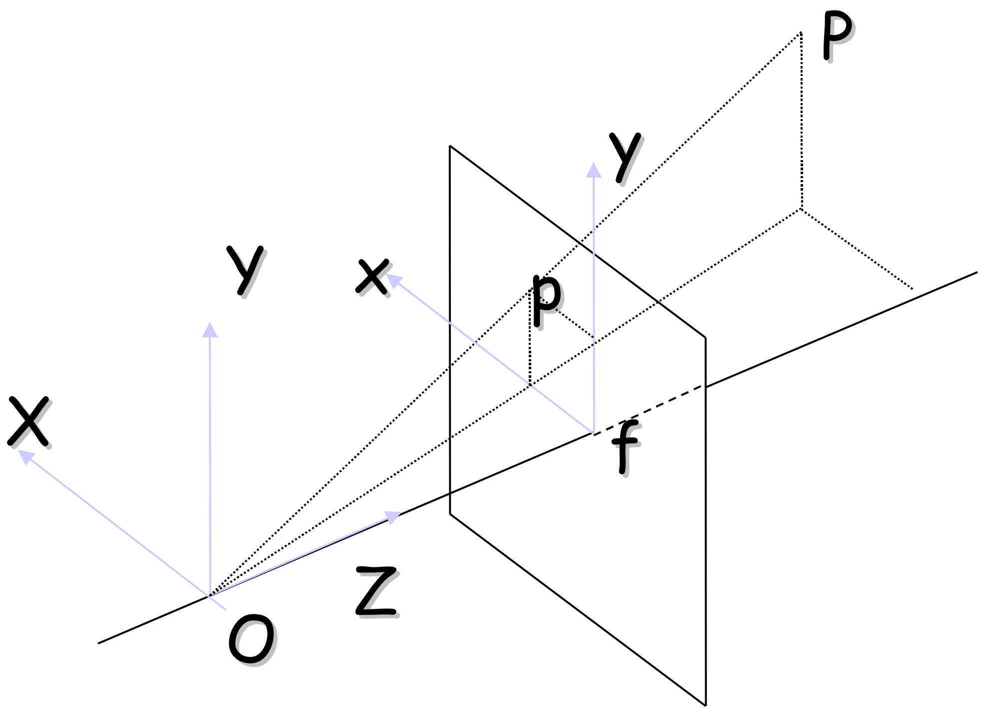


$$\frac{x}{f} = \frac{X}{Z}$$

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

Adapted from Octavia Camps

Pinhole Camera Model



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

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

Adapted from Octavia Camps

Perspective Matrix Equation

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

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

Using homogeneous coordinates:

$$\begin{bmatrix} x' \\ y' \\ z' \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$x = \frac{x'}{z'} \qquad y = \frac{y'}{z'}$$

Adapted from Octavia Camps

Perspective Matrix Equation

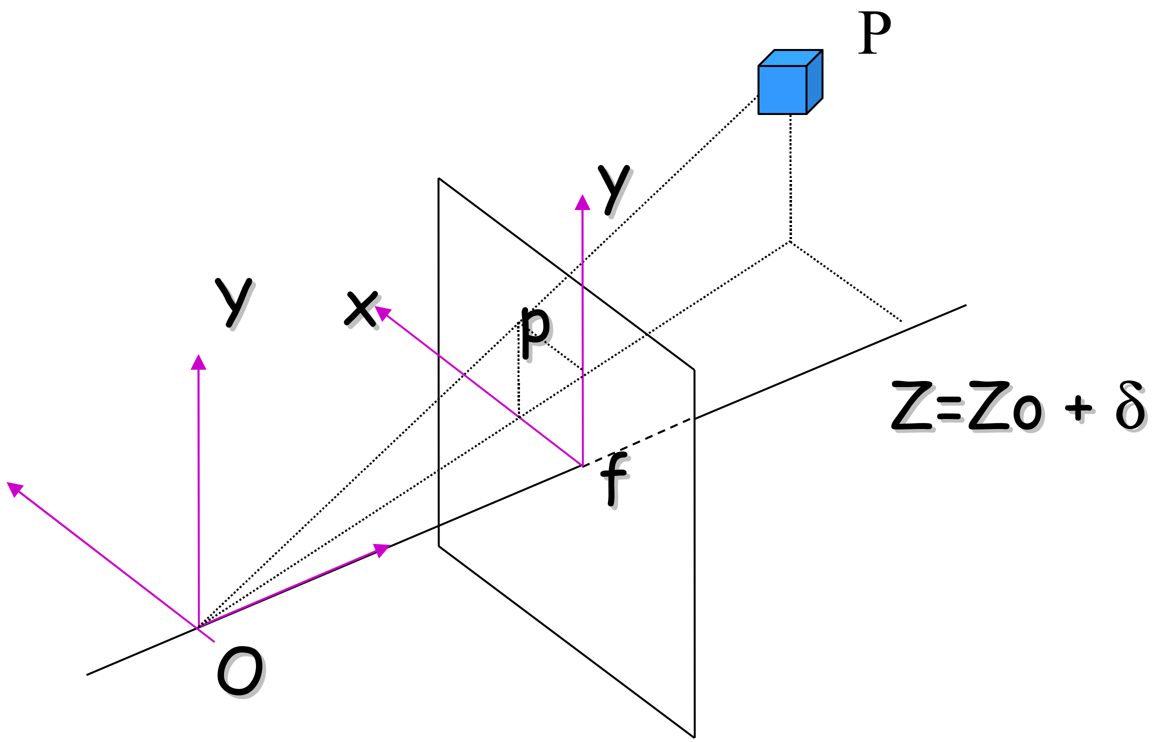
- Homogenous coordinates for 3D
  - four coordinates for 3D point
  - equivalence relation (X,Y,Z,T) is the same as (k X, k Y, k Z,k T)
- Turn previous expression into HC's
  - HC's for 3D point are (X,Y,Z,T)
  - HC's for point in image are (U,V,W)

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{f} & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

$$(U,V,W) \rightarrow (\frac{U}{W}, \frac{V}{W}) = (u,v)$$

Adapted from Gregory Hager, JHU

Weak Perspective Model

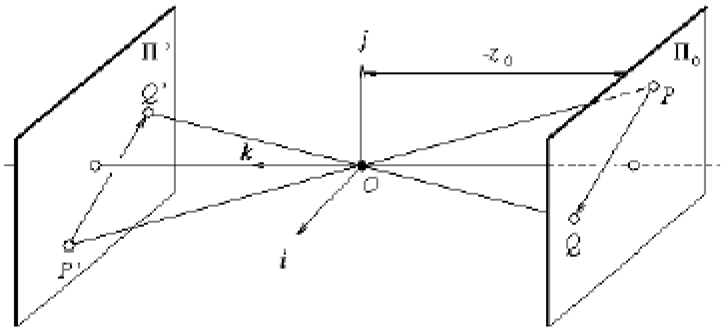


$$x = f X / Zo$$
$$y = f Y / Zo$$

- Object depth δ << Camera distance Zo
- Linear equations !!

Adapted from Octavia Camps, PennState

Model for Weak Perspective Projection



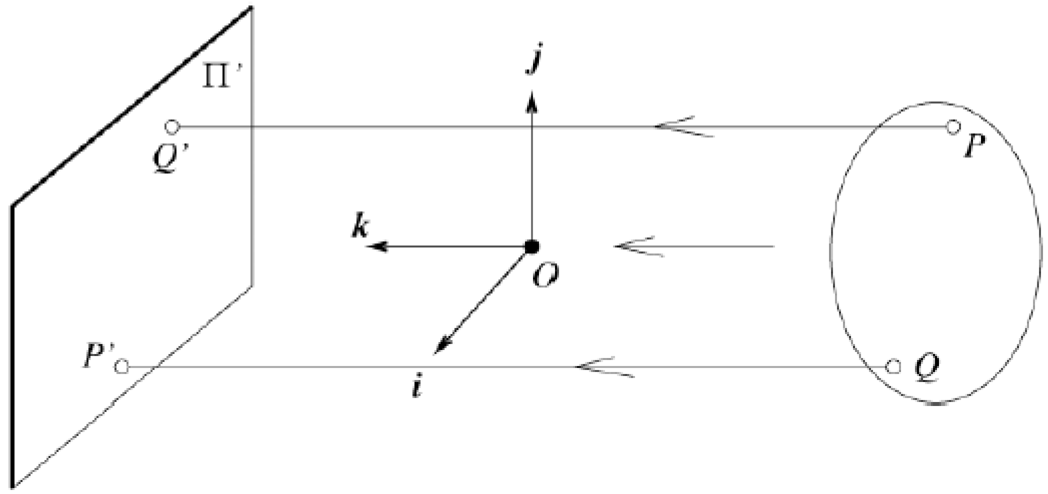
$$u = sx$$

$$v = sy$$

$$s = f / Z^*$$

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & Z^* / f \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Orthographic Projection



Suppose I let f go to infinity; then

$$u = x$$

$$v = y$$

Forsyth & Ponce

The projection matrix for orthographic projection

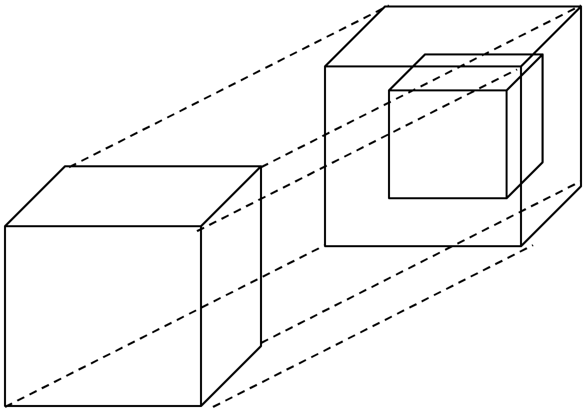
$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

$$= \begin{pmatrix} X \\ Y \\ T \end{pmatrix} \rightarrow \frac{1}{T} \begin{pmatrix} X \\ Y \end{pmatrix}$$

HC Non-HC

Adapted from David Forsyth, UC Berkeley

Weak Perspective vs Ortographic Projection



Weak perspective = Orthographic projection +
Isotropic Scaling

Adapted from Octavia Camps, PennState

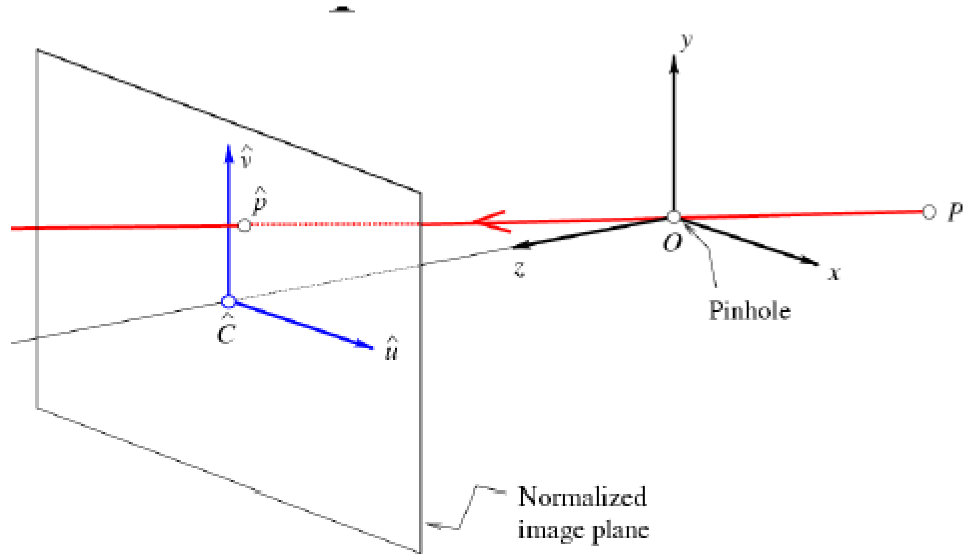
Camera parameters

- Intrinsic parameters
  - Focal length, principal point, aspect ratio, angle between axes
- Extrinsic parameters
  - Translation, and Rotation parameters

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{intrinsic parameters} \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} \text{Transformation} \\ \text{representing} \\ \text{extrinsic parameters} \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

Adapted from David Forsyth, UC Berkeley

Intrinsic parameters



Forsyth&Ponce

Perspective projection

$$u = f\frac{x}{z}$$

$$v = f\frac{y}{z}$$

Adapted from Trevor Darrell, MIT

Intrinsic parameters – focal length

$$\begin{pmatrix} U \\ V \\ W \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \\ T \end{pmatrix}$$

$$(U, V, W) \rightarrow (\frac{U}{W}, \frac{V}{W}) = (u, v)$$

$$p = M_{int} \,.\, P$$

Adapted from Octavia Camps, PennState

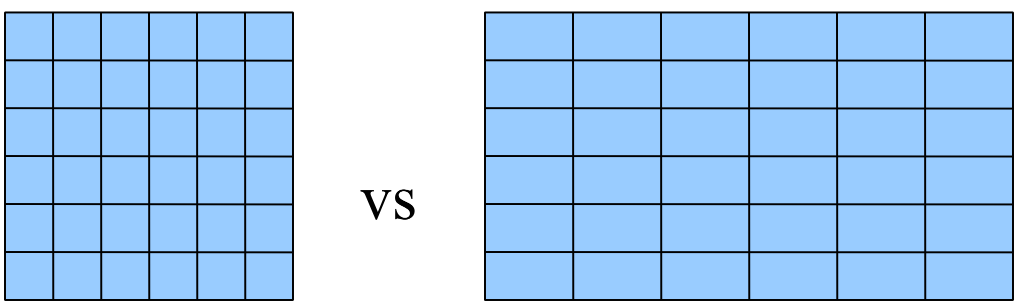
Intrinsic parameters – aspect ratio

- The CCD sensor is made of a rectangular grid nxm of photosensors.
- Each photosensor generates an analog signal that is digitized by a frame grabber into an array of NxM pixels.

$$u = \alpha \frac{x}{z}$$

Pixels may not be square

$$v = \beta \frac{y}{z}$$



$$M_{int} = \begin{bmatrix} \alpha & 0 & 0 & 0 \\ 0 & \beta & 0 & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

Adapted from Octavia Camps, PennState

Intrinsic parameters - origin

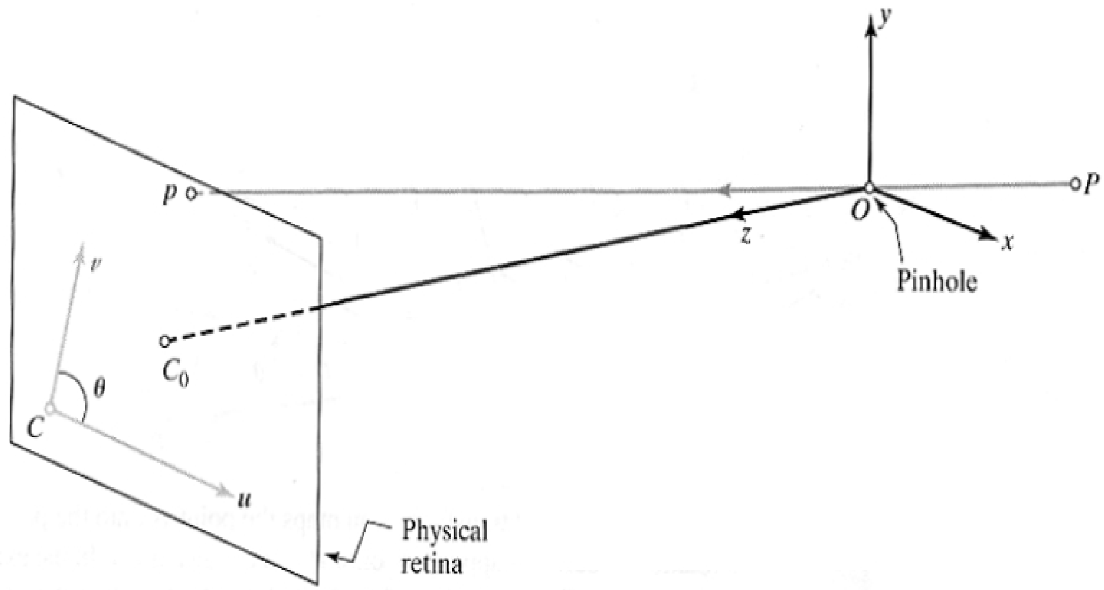
We don't know the origin of our camera pixel coordinates

$$u = \alpha \frac{x}{z} + u_0$$

$$v = \beta \frac{y}{z} + v_0$$

$$M_{int} = \begin{bmatrix} \alpha & 0 & uo & 0 \\ 0 & \beta & vo & 0 \\ 0 & 0 & 1/f & 0 \end{bmatrix}$$

Intrinsic parameters – angle between axes

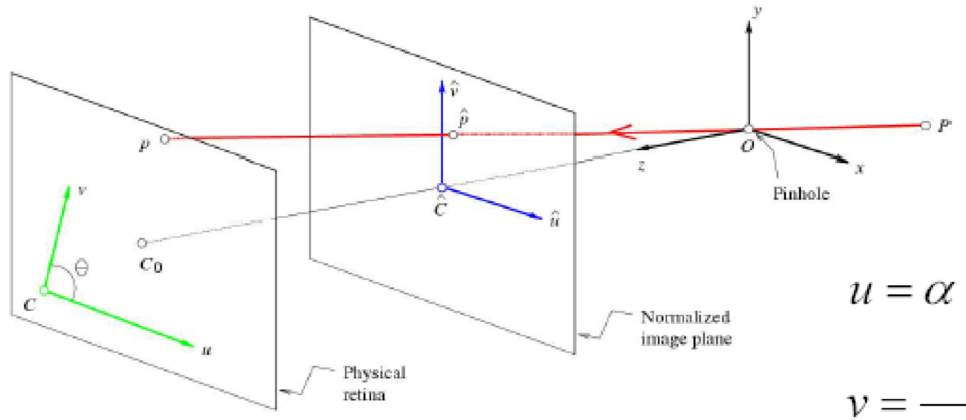


May be skew between camera pixel axes

$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Intrinsic parameters



$$u = \alpha \frac{x}{z} - \alpha \cot(\theta) \frac{y}{z} + u_0$$

$$v = \frac{\beta}{\sin(\theta)} \frac{y}{z} + v_0$$

Using homogenous coordinates, we can write this as:

$$\begin{pmatrix} u \\ v \\ 1 \end{pmatrix} = \frac{1}{z} \begin{pmatrix} \alpha & -\alpha \cot(\theta) & u_0 & 0 \\ 0 & \frac{\beta}{\sin(\theta)} & v_0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \\ 1 \end{pmatrix}$$

or:

$$\vec{p} = \frac{1}{z} \begin{pmatrix} K & \vec{0} \end{pmatrix} \vec{P}$$

Adapted from Trevor Darrell, MIT

Extrinsic parameters

Translation and rotation of camera frame

$$ {}^{C}P = {}^{C}_{W}R \, {}^{W}P + {}^{C}O_{W} $$

Non-homogeneous coordinates

$$ \begin{pmatrix} C_X \\ C_Y \\ C_Z \\ 1 \end{pmatrix} = \begin{pmatrix} - & - & - & | \\ - & {}^{C}_{W}R & - & {}^{C}O_{W} \\ - & - & - & | \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} W_X \\ W_Y \\ W_Z \\ 1 \end{pmatrix} $$

Homogeneous coordinates

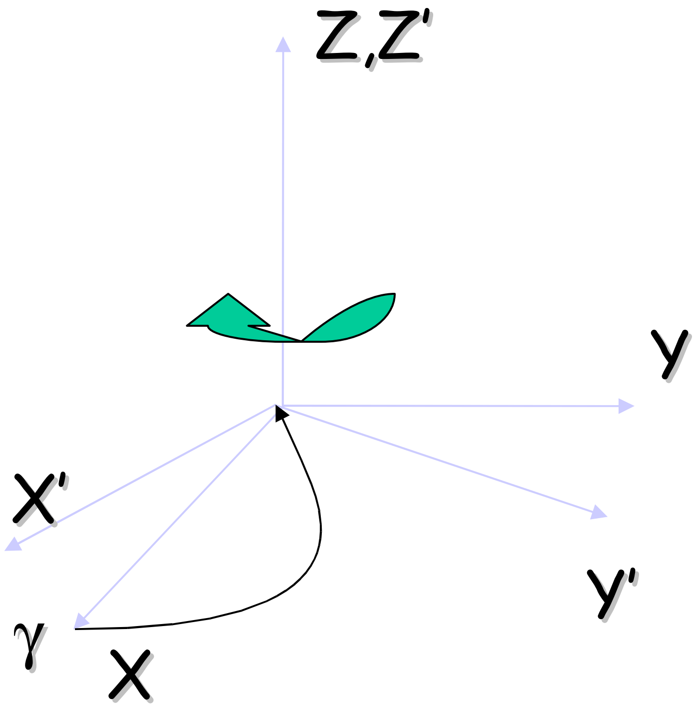
$$ \begin{pmatrix} {}^{C}P \\ 1 \end{pmatrix} = \begin{pmatrix} {}^{C}_{W}\mathcal{R} & {}^{C}O_{W} \\ \mathbf{0}^{T} & 1 \end{pmatrix} \begin{pmatrix} {}^{W}P \\ 1 \end{pmatrix} $$

Block matrix form

Adapted from Trevor Darrell, MIT

3D Rotation of Coordinates Systems

Rotation around the coordinate axes, **clock-clockwise**:



$$R_x(\alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha \\ 0 & \sin\alpha & \cos\alpha \end{bmatrix}$$

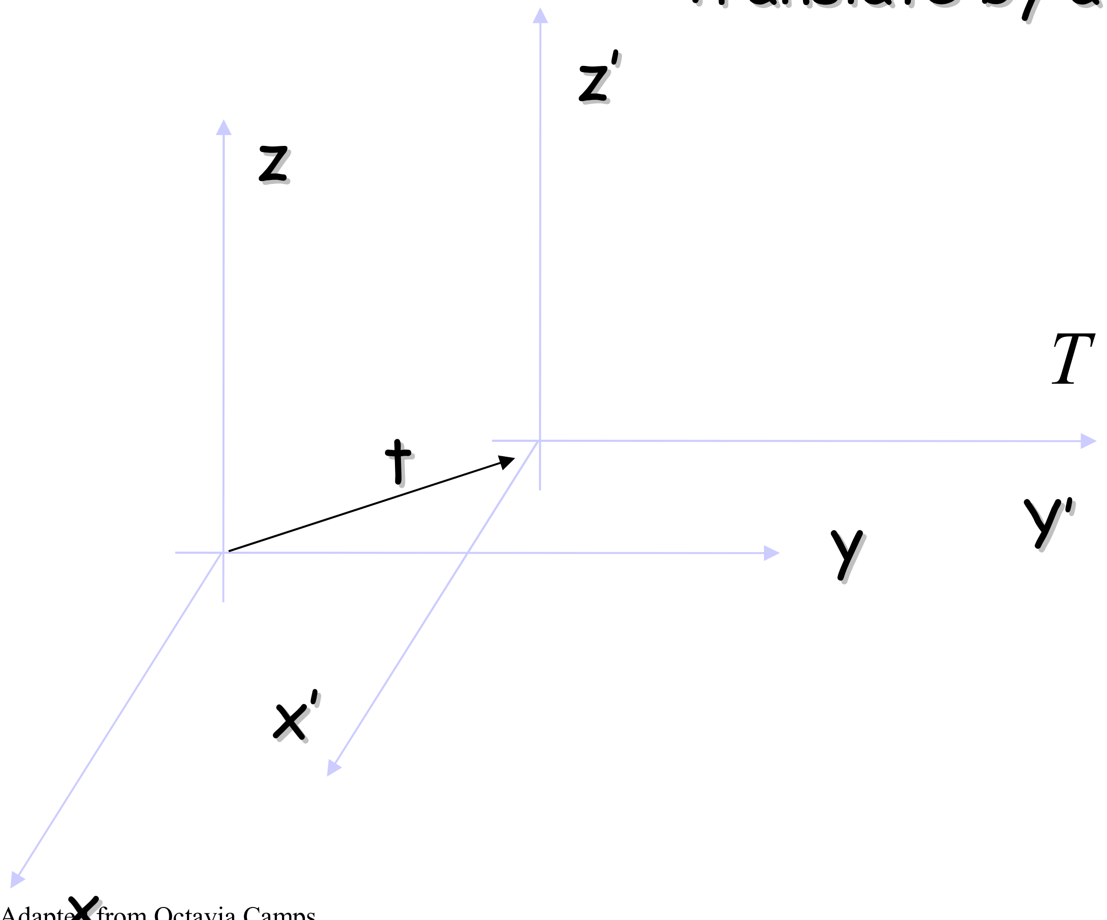
$$R_y(\beta) = \begin{bmatrix} \cos\beta & 0 & \sin\beta \\ 0 & 1 & 0 \\ -\sin\beta & 0 & \cos\beta \end{bmatrix}$$

$$R_z(\gamma) = \begin{bmatrix} \cos\gamma & -\sin\gamma & 0 \\ \sin\gamma & \cos\gamma & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Adapted from Octavia Camps

3D Translation of Coordinate Systems

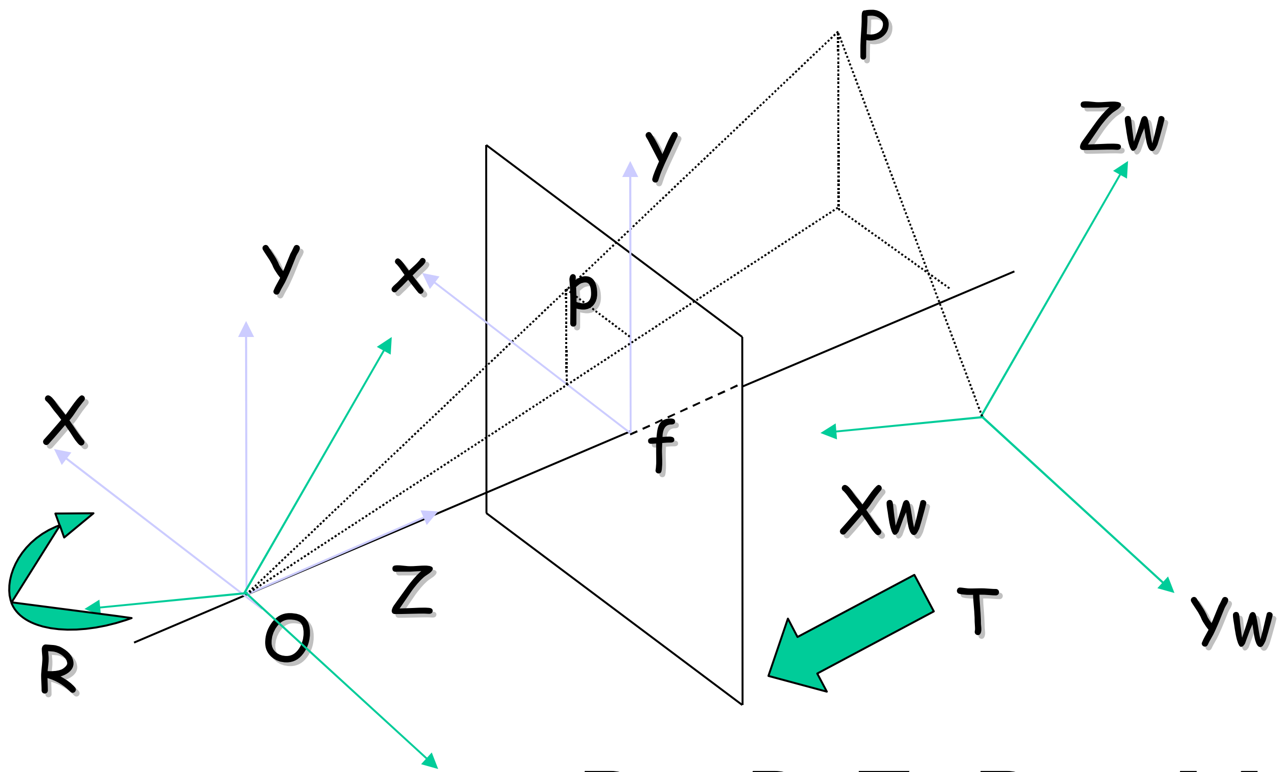
Translate by a vector $t=(t_x,t_y,t_x)^T$:



$$T = \begin{bmatrix} 1 & 0 & 0 & -t_x \\ 0 & 1 & 0 & -t_y \\ 0 & 0 & 1 & -t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Adapted from Octavia Camps

Combining Extrinsic and Intrinsic Parameters



$$P = R \cdot T \cdot P_w = M_{ext} \cdot P_w$$

$$p = M_{int} P = M_{int} M_{ext} \cdot P_w$$

Adapted from Octavia Camps

Combining Extrinsic and Intrinsic parameters

$$\vec{p} = \frac{1}{z}\begin{pmatrix} K & \vec{0} \end{pmatrix} \vec{P}$$

Intrinsic

$${}^{C}P = {}^{C}_{W}R\ {}^{W}P + {}^{C}O_{W}$$

Extrinsic

$$\vec{p} = \frac{1}{z} K \begin{pmatrix} {}^{C}_{W}R & {}^{C}O_{W} \end{pmatrix} \vec{P}$$

$$\vec{p} = \frac{1}{z} M \vec{P}$$

Forsyth & Ponce

Combining Extrinsic and Intrinsic parameters

$$\boldsymbol{p} = \frac{1}{z}\mathcal{M}\boldsymbol{P}, \quad \text{where} \quad \mathcal{M} = \mathcal{K}(\mathcal{R} \quad \boldsymbol{t}), \tag{2.15}$$

$\mathcal{R} = {}^C_W\mathcal{R}$ is a rotation matrix, $\boldsymbol{t} = {}^C O_W$ is a translation vector, and $\boldsymbol{P} = ({}^W x, {}^W y, {}^W z, 1)^T$ denotes the *homogeneous* coordinate vector of P in the frame (W).

A projection matrix can be written explicitly as a function of its five intrinsic parameters (α, β, u_0, v_0, and θ) and its six extrinsic ones (the three angles defining $\mathcal{R}$ and the three coordinates of $\boldsymbol{t}$), namely,

$$\mathcal{M} = \begin{pmatrix} \alpha \boldsymbol{r}_1^T - \alpha \cot\theta \boldsymbol{r}_2^T + u_0 \boldsymbol{r}_3^T & \alpha t_x - \alpha \cot\theta t_y + u_0 t_z \\ \frac{\beta}{\sin\theta}\boldsymbol{r}_2^T + v_0 \boldsymbol{r}_3^T & \frac{\beta}{\sin\theta} t_y + v_0 t_z \\ \boldsymbol{r}_3^T & t_z \end{pmatrix}, \tag{2.17}$$

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

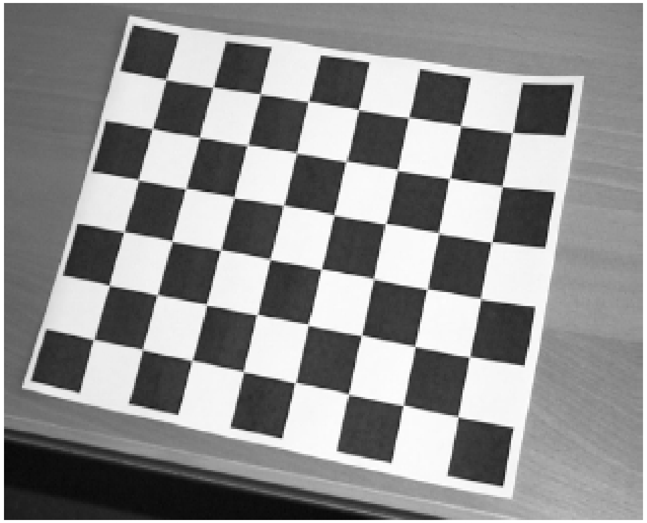
Forsyth & Ponce

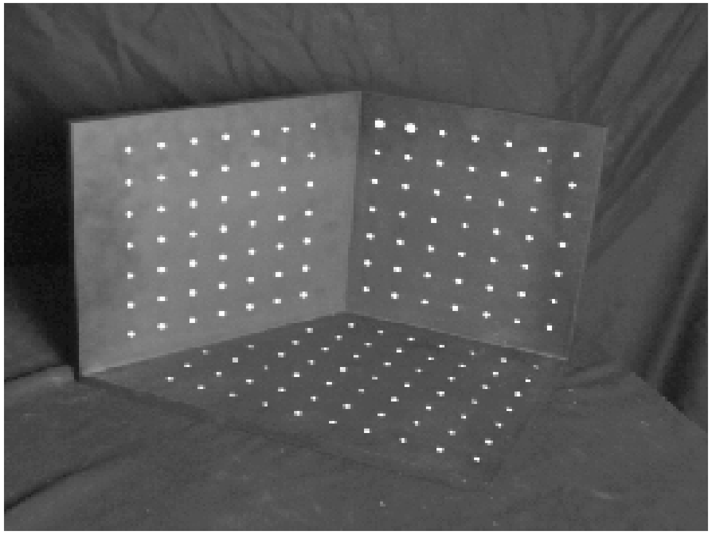
Camera Calibration

Compute the camera intrinsic and extrinsic parameters using only observed camera data

Place a known object in the scene

- identify correspondence between image and scene
- compute mapping from scene to image





Camera Calibration

$$\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix} \cong \begin{bmatrix} m_{00} & m_{01} & m_{02} & m_{03} \\ m_{10} & m_{11} & m_{12} & m_{13} \\ m_{20} & m_{21} & m_{22} & m_{23} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \\ Z_i \\ 1 \end{bmatrix}$$

$$u_i = \frac{m_{00}X_i + m_{01}Y_i + m_{02}Z_i + m_{03}}{m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}}$$

$$v_i = \frac{m_{10}X_i + m_{11}Y_i + m_{12}Z_i + m_{13}}{m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}}$$

Adapted from Trevor Darrell

Camera Calibration

$$u_i(m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}) = m_{00}X_i + m_{01}Y_i + m_{02}Z_i + m_{03}$$
$$v_i(m_{20}X_i + m_{21}Y_i + m_{22}Z_i + m_{23}) = m_{10}X_i + m_{11}Y_i + m_{12}Z_i + m_{13}$$

$$\begin{bmatrix} X_i & Y_i & Z_i & 1 & 0 & 0 & 0 & 0 & -u_iX_i & -u_iY_i & -u_iZ_i & -u_i \\ 0 & 0 & 0 & 0 & X_i & Y_i & Z_i & 1 & -v_iX_i & -v_iY_i & -v_iZ_i & -v_i \end{bmatrix} \begin{bmatrix} m_{00} \\ m_{01} \\ m_{02} \\ m_{03} \\ m_{10} \\ m_{11} \\ m_{12} \\ m_{13} \\ m_{20} \\ m_{21} \\ m_{22} \\ m_{23} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Adapted from Trevor Darrell

Camera Calibration

$$\begin{bmatrix} X_1 & Y_1 & Z_1 & 1 & 0 & 0 & 0 & 0 & -u_1X_1 & -u_1Y_1 & -u_1Z_1 & -u_1 \\ 0 & 0 & 0 & 0 & X_1 & Y_1 & Z_1 & 1 & -v_1X_1 & -v_1Y_1 & -v_1Z_1 & -v_1 \\ & & & & & & \vdots & & & & & \\ X_n & Y_n & Z_n & 1 & 0 & 0 & 0 & 0 & -u_nX_n & -u_nY_n & -u_nZ_n & -u_n \\ 0 & 0 & 0 & 0 & X_n & Y_n & Z_n & 1 & -v_nX_n & -v_nY_n & -v_nZ_n & -v_n \end{bmatrix} \begin{bmatrix} m_{00} \\ m_{01} \\ m_{02} \\ m_{03} \\ m_{10} \\ m_{11} \\ m_{12} \\ m_{13} \\ m_{20} \\ m_{21} \\ m_{22} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{bmatrix}$$

M has 12 entries
each image point provides 2 equations
Can solve m_{ij}'s by Least Square Solution

Adapted from Trevor Darrell