Geometric Camera Parameters

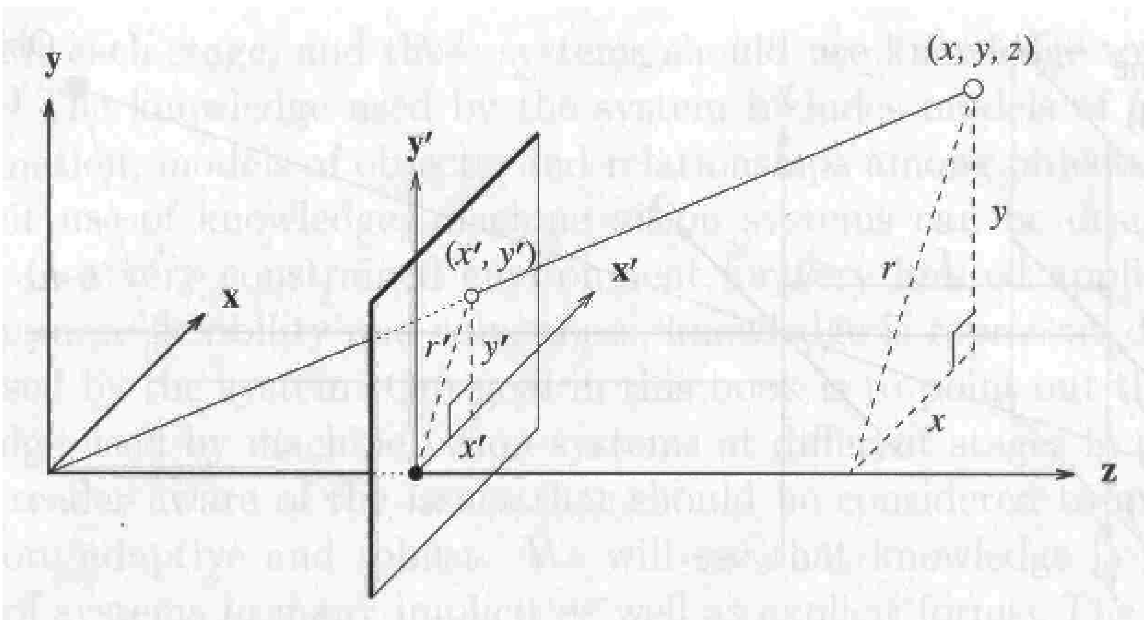
• What assumptions have we made so far?

- All equations we have derived for far are written in the camera reference frames.

- These equations are valid only when:

  (1) all distances are measured in the camera's reference frame.

  (2) the image coordinates have their origin at the principal point.



- In general, the world and pixel coordinate systems are related by a set of physical parameters such as:

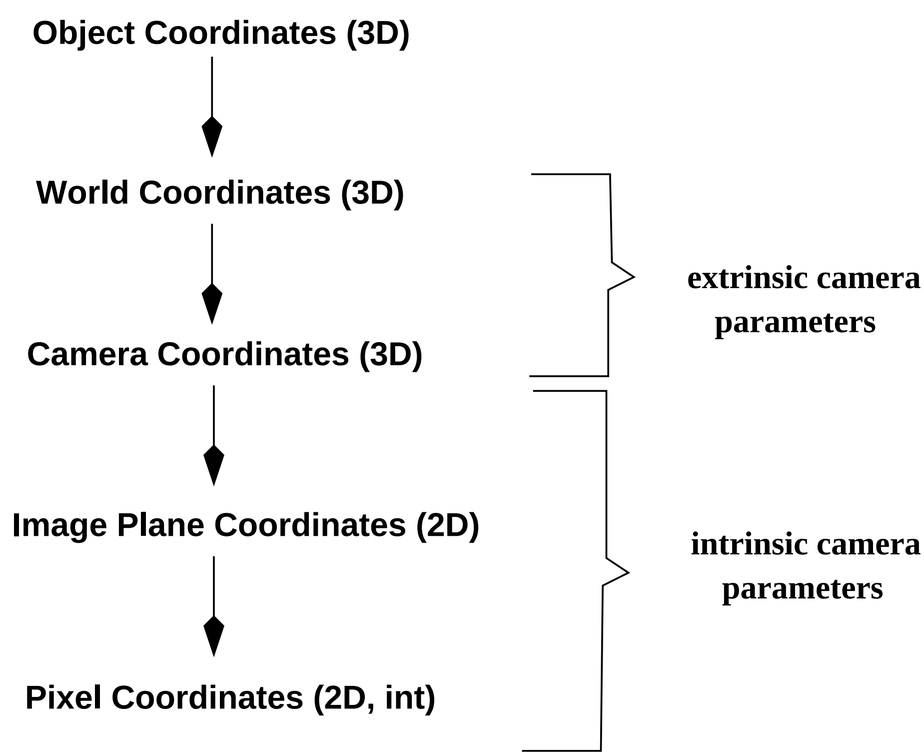
  * the focal length of the lens
  * the size of the pixels
  * the position of the principal point
  * the position and orientation of the camera

• Types of parameters (Trucco 2.4)

- Two types of parameters need to be recovered in order for us to reconstruct the 3D structure of a scene from the pixel coordinates of its image points:

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

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



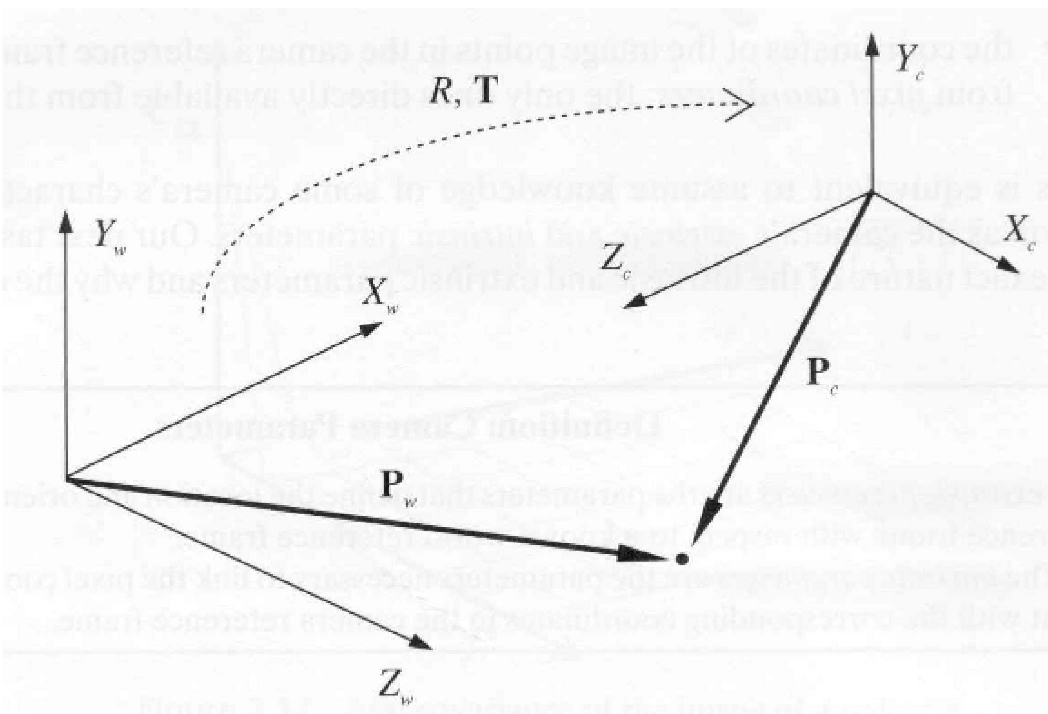
• Extrinsic camera parameters

- These are the parameters that identify uniquely the transformation between the *unknown camera reference frame* and the *known world reference frame*.

- Typically, determining these parameters means:

(1) finding the translation vector between the relative positions of the origins of the two reference frames.

(2) finding the rotation matrix that brings the corresponding axes of the two frames into alignment (i.e., onto each other)



- Using the extrinsic camera parameters, we can find the relation between the coordinates of a point P in world (P_w) and camera (P_c) coordinates:

$$P_c = R(P_w - T) \text{ where } R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

- If $P_c = \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$ and $P_w = \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix}$, then

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w - T_x \\ Y_w - T_y \\ Z_w - T_z \end{bmatrix}$$

or

$$X_c = R_1^T (P_w - T)$$

$$Y_c = R_2^T (P_w - T)$$

$$Z_c = R_3^T (P_w - T)$$

where R_i^T corresponds to the i-th row of the rotation matrix

• Intrinsic camera parameters

- These are the parameters that characterize the optical, geometric, and digital characteristics of the camera:

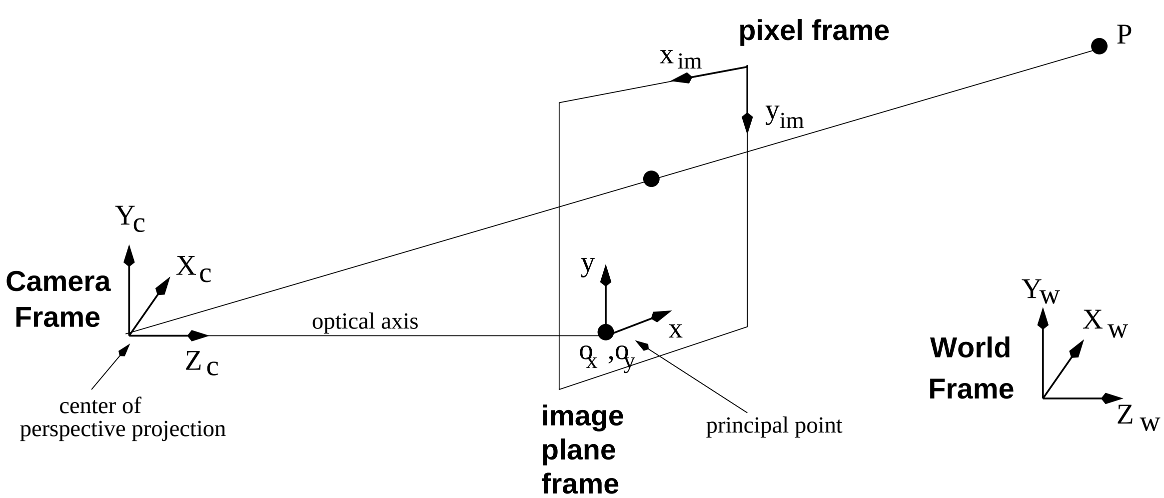
(1) the perspective projection (focal length f).
(2) the transformation between image plane coordinates and pixel coordinates.
(3) the geometric distortion introduced by the optics.

From Camera Coordinates to Image Plane Coordinates

- Apply perspective projection:

$$x = f\,\frac{X_c}{Z_c} = f\,\frac{R_1^T(P_w - T)}{R_3^T(P_w - T)}, \qquad y = f\,\frac{Y_c}{Z_c} = f\,\frac{R_2^T(P_w - T)}{R_3^T(P_w - T)}$$

From Image Plane Coordinates to Pixel coordinates



$$x = -(x_{im} - o_x)s_x \quad \text{or} \quad x_{im} = -x/s_x + o_x$$

$$y = -(y_{im} - o_y)s_y \quad \text{or} \quad y_{im} = -y/s_y + o_y$$

where (o_x, o_y) are the coordinates of the principal point (in pixels, e.g., $o_x = N/2$, $o_y = M/2$ if the principal point is the center of the image) and s_x, s_y correspond to the effective size of the pixels in the horizontal and vertical directions (in millimeters).

- Using matrix notation:

$$\begin{bmatrix} x_{im} \\ y_{im} \\ 1 \end{bmatrix} = \begin{bmatrix} -1/s_x & 0 & o_x \\ 0 & -1/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

Relating pixel coordinates to world coordinates

$$-(x_{im} - o_x)s_x = f \frac{R_1^T(P_w - T)}{R_3^T(P_w - T)}, \quad -(y_{im} - o_y)s_y = f \frac{R_2^T(P_w - T)}{R_3^T(P_w - T)}$$

or

$$x_{im} = -fs_x \frac{R_1^T(P_w - T)}{R_3^T(P_w - T)} + o_x, \quad y_{im} = -fs_y \frac{R_2^T(P_w - T)}{R_3^T(P_w - T)} + o_y$$

Image distortions due to optics

Assuming *radial* distortion:

$$x = x_d(1 + k_1 r^2 + k_2 r^4)$$

$$y = y_d(1 + k_1 r^2 + k_2 r^4)$$

where (x_d, y_d) are the coordinates of the distorted points ($r^2 = x_d^2 + y_d^2$)

k_1 and k_2 are intrinsic parameters too but will not be considered here...

• Combine extrinsic with intrinsic camera parameters

- The matrix containing the intrinsic camera parameters:

$$M_{in} = \begin{bmatrix} -f/s_x & 0 & o_x \\ 0 & -f/s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}$$

- The matrix containing the extrinsic camera parameters:

$$M_{ex} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & -R_1^T T \\ r_{21} & r_{22} & r_{23} & -R_2^T T \\ r_{31} & r_{32} & r_{33} & -R_3^T T \end{bmatrix}$$

- Using homogeneous coordinates:

$$\begin{bmatrix} x_h \\ y_h \\ w \end{bmatrix} = M_{in}\ M_{ex} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = M \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} m_{11} & m_{12} & m_{13} & m_{14} \\ m_{21} & m_{22} & m_{23} & m_{24} \\ m_{31} & m_{32} & m_{33} & m_{34} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

- Homogenization is needed to obtain the pixel coordinates:

$$x_{im} = \frac{x_h}{w} = \frac{m_{11}X_w + m_{12}Y_w + m_{13}Z_w + m_{14}}{m_{31}X_w + m_{32}Y_w + m_{33}Z_w + m_{34}}$$

$$y_{im} = \frac{y_h}{w} = \frac{m_{21}X_w + m_{22}Y_w + m_{23}Z_w + m_{24}}{m_{31}X_w + m_{32}Y_w + m_{33}Z_w + m_{34}}$$

- M is called the *projection matrix* (it is a 3 x 4 matrix).

Note: the relation of 3D points and their 2D projections can be seen as a linear transformation from the projective space $(X_w, Y_w, Z_w, 1)^T$ to the projective plane $(x_h, y_h, w)^T$.

• The perspective camera model (using matrix notation)

- Assuming $o_x = o_y = 0$ and $s_x = s_y = 1$

$$M_p = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & fR_1^T T \\ -fr_{21} & -fr_{22} & -fr_{23} & fR_2^T T \\ r_{31} & r_{32} & r_{33} & -R_3^T T \end{bmatrix}$$

- Let's verify the correctness of the above matrix:

$$p = M_p P_w = \begin{bmatrix} -fR_1^T & fR_1^T T \\ -fR_2^T & fR_2^T T \\ R_3^T & -R_3^T T \end{bmatrix} \begin{bmatrix} P_w \\ 1 \end{bmatrix} = \begin{bmatrix} -fR_1^T(P_w - T) \\ -fR_2^T(P_w - T) \\ R_3^T(P_w - T) \end{bmatrix}$$

- After homogenization (we get the same equations as in page 23):

$$x = -f \frac{R_1^T(P_w - T)}{R_3^T(P_w - T)} \qquad y = -f \frac{R_2^T(P_w - T)}{R_3^T(P_w - T)}$$

• The weak perspective camera model (using matrix notation)

$$M_{wp} = \begin{bmatrix} -fr_{11} & -fr_{12} & -fr_{13} & fR_1^T T \\ -fr_{21} & -fr_{22} & -fr_{23} & fR_2^T T \\ 0 & 0 & 0 & R_3^T(\bar{P} - T) \end{bmatrix}$$

where $\bar{P}$ is the centroid of the object (i.e., object's average distance from the camera)

- We can verify the correctness of the above matrix:

$$p = M_{wp} P_w = \begin{bmatrix} -fR_1^T & fR_1^T T \\ -fR_2^T & fR_2^T T \\ 0 \;\; 0 \;\; 0 & R_3^T(\bar{P} - T) \end{bmatrix} \begin{bmatrix} P_w \\ 1 \end{bmatrix} = \begin{bmatrix} -fR_1^T(P_w - T) \\ -fR_2^T(P_w - T) \\ R_3^T(\bar{P} - T) \end{bmatrix}$$

- After homogenization:

$$x = -f \frac{R_1^T(P_w - T)}{R_3^T(\bar{P} - T)} \qquad y = -f \frac{R_2^T(P_w - T)}{R_3^T(\bar{P} - T)}$$

• The affine camera model

- The entries of the projection matrix are totally unconstrained:

$$M_a = \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ 0 & 0 & 0 & a_{34} \end{bmatrix}$$

- The affine model does not appear to correspond to any physical camera.

- Leads to simple equations and appealing geometric properties.

- Does not preserve angles but does preserve parallelism.