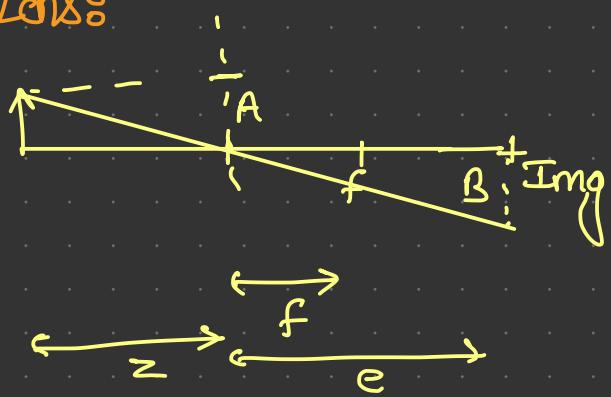


Pinhole Cameras

- Barrier reduces blurring
- Opening is called aperture

↳ More aperture, more blurring
 ↳ If too small, less light gets through
 diffraction effects

Thin Lens



$$\frac{B}{A} = \frac{e}{z}, \frac{e}{f} = 1 + \frac{e}{z} \Rightarrow \frac{1}{f} = \frac{1}{e} + \frac{1}{z}$$

→ Specific dist for "in focus"
 ↳ Length of lens/pinhole

$$\rightarrow \text{Blur Circle: } \frac{\angle \delta}{2e} = R$$

- If obj too far ($z \gg f$),
 img formed at f .
- Perspective: Dependence of apparent size of object on depth (z)
 - Far objects are smaller
 - Can perceive depth

→ Minimal $\angle \delta \Rightarrow$ Minimal R
 → R is smaller than image resolution
 for good image

- Projective:
 - Straight lines are still straight
 - Loses length & angles

- Vanishing points:
 Int. of 11 lines

- Depth of field [DoF]
 - Dist b/w nearest and farthest objects that are acceptably sharp

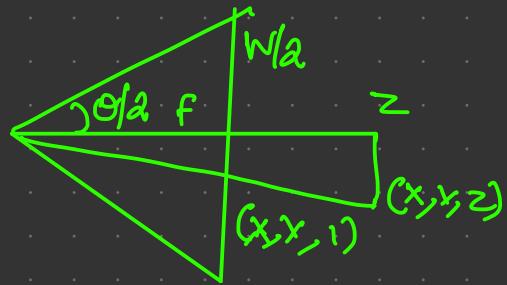
→ Effect of aperture on DoF

Aperature $\downarrow \Rightarrow$ ↑ objects in focus but reduces amount of light

Field of view

$\uparrow f \Rightarrow$ Wider field of view [more world points]
 $\downarrow f \Rightarrow$ Narrow field of view [less world points]

Small FOV \Rightarrow ↑ focus



$$\tan\left(\frac{\theta}{2}\right) = \frac{w}{2f} \Rightarrow f = \frac{w}{2} \left[\tan\left(\frac{\theta}{2}\right) \right]^{-1}$$

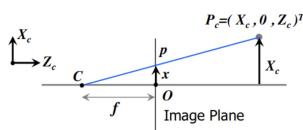
Perspective Camera:

→ Camera measures angles & not distances

→ Camera plane → Image Planes

Perspective Projection (1)

From the Camera frame to the image plane



The Camera point $P_c = (X_c, 0, Z_c)^T$ projects to $p = (x, y)$ onto the image plane

From similar triangles: $\frac{x}{f} = \frac{X_c}{Z_c} \Rightarrow x = \frac{f X_c}{Z_c}$

Similarly, in the general case:

$$\frac{y}{f} = \frac{Y_c}{Z_c} \Rightarrow y = \frac{f Y_c}{Z_c}$$

1. Convert P_c to image-plane coordinates (x, y)

2. Convert P_c to (discretised) pixel coordinates (u, v)

From World to Pixel coordinates

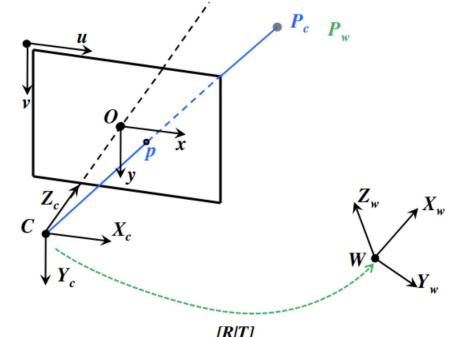
Find pixel coordinates (u, v) of point P_w in the world frame:

0. Convert world point P_w to camera point P_c

Find pixel coordinates (u, v) of point P_c in the camera frame:

1. Convert P_c to image-plane coordinates (x, y)

2. Convert P_c to (discretised) pixel coordinates (u, v)



$$x = \frac{f X_c}{Z_c}, \quad y = \frac{f Y_c}{Z_c}$$

k_u, k_v are scale factors taking f in CMS to pixels

$k_u, k_v = \frac{\text{Image Width/Height}}{\text{Sensor width/height}}$

Perspective Projection (2)

From the Camera frame to pixel coordinates

To convert p from the local image plane coords (x, y) to the pixel coords (u, v) , we need to account for:

- the pixel coords of the camera optical center $O = (u_0, v_0)$
- Scale factors k_u, k_v for the pixel-size in both dimensions

$$\text{So: } u = u_0 + k_u x \Rightarrow u = u_0 + \frac{k_u f X_c}{Z_c}$$

$$v = v_0 + k_v y \Rightarrow v = v_0 + \frac{k_v f Y_c}{Z_c}$$

Use Homogeneous Coordinates for linear mapping from 3D to 2D, by introducing an extra element (scale):

$$p = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix} \rightarrow \tilde{p} = \begin{pmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \\ 1 \end{pmatrix} = \lambda \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

Perspective Projection (3)

So:

$$u = u_0 + \frac{k_u f X_c}{Z_c}$$

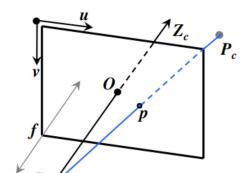
$$v = v_0 + \frac{k_v f Y_c}{Z_c}$$

Expressed in matrix form and homogeneous coordinates:

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} k_u f & 0 & u_0 \\ 0 & k_v f & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$

Or alternatively

$$\begin{bmatrix} \lambda u \\ \lambda v \\ \lambda \end{bmatrix} = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = K \begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix}$$



Focal length in pixels

K is called "Calibration matrix" or "Matrix of Intrinsic Parameters"

Sometimes, it is common to assume a skew factor ($K_{12} \neq 0$) to account for possible misalignments between CCD and lens. However, the camera manufacturing process today is so good that we can safely assume $K_{12} = 0$ and $\alpha_u = \alpha_v$.

$$K = \begin{bmatrix} \alpha_u & 0 & u_0 \\ 0 & \alpha_v & v_0 \\ 0 & 0 & 1 \end{bmatrix}$$

is calibration matrix
 Matrix of intrinsic parameters
 where α_u, α_v are focal length in pixels

→ Proof of vanishing point can be showed by taking above equation

Perspective Projection (4)

From the Camera frame to the World frame

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \end{bmatrix} + \begin{bmatrix} t_1 \\ t_2 \\ t_3 \end{bmatrix}$$

$$\begin{bmatrix} X_c \\ Y_c \\ Z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix} = \begin{bmatrix} R & T \\ 1 & 1 \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

K = Scaling factor that takes from image to pixel coordinate

Assumptions:
 Image coordinates are in center of image

Extrinsic Parameters: $[R|T]$

Central projection: $\vec{x} \approx K\vec{x}$

Perspective Projection Equation:

$$\lambda \vec{x}_c = K \vec{x} \quad \text{Extrinsic Parameters}$$

$$\vec{x} \approx K\vec{x} \quad \text{Central projection}$$

IMP →

λ = Depth ($\lambda = 2c$) of scene point

$$\vec{x} \approx K\vec{x} \Rightarrow K^{-1}\lambda\vec{x} = \vec{x}$$

$$\lambda K^{-1}\vec{x} = \vec{x} \quad \text{i.e. } K^{-1}\vec{x} \text{ is parallel to } \vec{x}$$

points to direction of 3D vector \vec{x}

$$\lambda \vec{x}_c = K T_w^c \vec{x}_w$$

$$= K [R | t] \vec{x}_w$$

↓
extrinsics → intrinsics

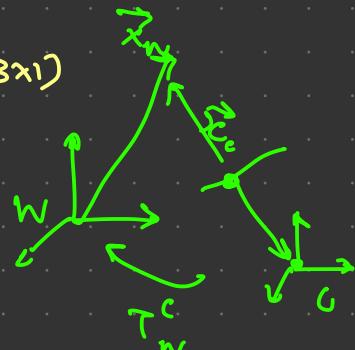
$$\vec{x}_c \approx K [R | t] \vec{x}_w \quad (3 \times 1)$$

$$\vec{x}_c = K [R | t] \vec{x}_w \quad (\text{as})$$

\vec{x}_c parallel to $K\vec{x}_c$

(as)

$$\text{cross}(\vec{x}_c, K\vec{x}_c) = 0$$



Distortions

- Trans. of ideal coordinates (u, v) to real observable coordinates
- Based on non-linear function

$$\text{Ex: } \begin{bmatrix} u_d \\ v_d \end{bmatrix} = (1 + k_2 \chi^2) \begin{bmatrix} u - u_0 \\ v - v_0 \end{bmatrix} + \begin{bmatrix} u_0 \\ v_0 \end{bmatrix}$$

$$\chi^2 = (u - u_0)^2 + (v - v_0)^2$$

Camera Calibration

Divide each v_i by v_{12} . or $P_{34} = 1$ and $y_{ij} = \frac{y_{ij}}{P_{34}}$
is the final solution for P .

How to get K from P ?

Before that let us look at $AP = 0_{2x1}$.
Since it is overdetermined and the observations are noisy \nexists NO P for which $AP = 0$.
Hence find the best P for which $\|AP\|^2$ is minimized or $P^T A^T A P$ is minimized such that $\|P\| = 1$.
The solution is by SVD.

$\vec{x}_c = [R_w^c \quad t_w^c] \vec{x}_w \rightarrow (1)$

$\vec{x}_c = K_{3x3} \vec{x}_w \rightarrow (2)$

$= K [R_w^c \quad t_w^c] \vec{x}_w \rightarrow (3) \Rightarrow P_{3x4} \vec{x}_w \rightarrow (4)$

$P_{3x4} \rightarrow \text{Camera Projection Matrix.}$

Camera Calibration: Process of estimating P .

Often (1) is rewritten as:

$$\vec{x}_c = R_w^c \vec{x}_w + t_w^c \rightarrow (5) = R_w^c \vec{x}_w - R_w^c t_w^c \rightarrow (6)$$

$$\text{or } \vec{x}_c = K R_w^c [\vec{x}_w - t_w^c] \rightarrow (6)$$

$$\text{or } \vec{x}_c = P_{3x4} \vec{x}_w \rightarrow (7)$$

$$\text{or rather } \vec{x}_c = P \vec{x}_w \text{ or } \vec{x}_c \approx P \vec{x}_w$$

$$\begin{aligned} \vec{x}_c &\approx P \vec{x}_w \\ \vec{x}_c &= P \vec{x}_w \\ \vec{x}_c &\approx K \vec{x}_c \\ \vec{x}_c &= K \vec{x}_c \end{aligned}$$

To estimate P :

↳ as a first step towards estimating K .

$P_{3x4} = 12$ parameters.

Every pair of correspondence $x_i \leftrightarrow \vec{x}_i$ gives two equations (Why?).

Hence 6 pairs of correspondences are needed to solve for P .

As a matter of fact one needs to solve for only 11 parameters of P as $\vec{x} \approx P \vec{x}$ is a homogeneous equation, which means

any $\gamma P \vec{x}$ also projects to same \vec{x} .

Hence every P_{ij} can be divided by P_{34} for example and P_{34} made 1.

Hence 6 pairs of correspondences $x_i \leftrightarrow \vec{x}_i$ are still required to solve for the 11 parameters of P .

How does one solve?

$x_i = P_{11} x_i + P_{12} y_i + P_{13} z_i + P_{14} \rightarrow (8)$

$y_i = P_{21} x_i + P_{22} y_i + P_{23} z_i + P_{24} \rightarrow (9)$

$z_i = P_{31} x_i + P_{32} y_i + P_{33} z_i + P_{34} \rightarrow (10)$

$x_i = \frac{x_i}{z_i}, y_i = \frac{y_i}{z_i} \rightarrow (11)$

as $\begin{bmatrix} x_i \\ y_i \\ z_i \end{bmatrix} = z_i \begin{bmatrix} x_i/z_i \\ y_i/z_i \\ 1 \end{bmatrix}$

Then $\vec{x}_i = \frac{P_{11} x_i + P_{12} y_i + P_{13} z_i + P_{14}}{P_{31} x_i + P_{32} y_i + P_{33} z_i + P_{34}}$

or $x_i P_{11} + y_i P_{12} + z_i P_{13} + P_{14} - P_{31} x_i x_i - P_{32} y_i x_i - P_{33} z_i x_i - P_{34} x_i = 0 \rightarrow (12)$

likewise $x_i P_{21} + y_i P_{22} + z_i P_{23} + P_{24} - P_{31} x_i y_i - P_{32} y_i y_i - P_{33} z_i y_i - P_{34} y_i = 0 \rightarrow (13)$

or $\begin{bmatrix} x_i & y_i & z_i & 1 & 0 & 0 & 0 & 0 & -x_i x_i & -y_i x_i & -z_i x_i & -x_i \\ 0 & 0 & 0 & 0 & x_i & y_i & z_i & 1 & -x_i y_i & -y_i y_i & -z_i y_i & -y_i \end{bmatrix} \begin{bmatrix} P_{11} \\ P_{12} \\ P_{13} \\ P_{14} \\ P_{21} \\ P_{22} \\ P_{23} \\ P_{24} \\ P_{31} \\ P_{32} \\ P_{33} \\ P_{34} \end{bmatrix} = 0 \cdot b_{12}$

↳ (15)

For every $\vec{x}_i \leftrightarrow \vec{x}_i$ we get 2 eqns of the form (15)

If there are $M > 6$ correspondences

We have $A_{2M \times 12} P = 0 \rightarrow (16)$

How to solve for P ?

Overdetermined set of equations

Avoid the trivial solution $P_{3x4} = 0_{3x4}$.

$SVD(A) = UDV^T$.

Last column of $V_{2M \times 12}$, a 12×1 column vector is the solution for P .

Divide each v_i by v_{12} . or $P_{34} = 1$ and $P_{ij} = \frac{P_{ij}}{P_{34}}$

is the final solution for P .