Camera Parameters:

$X_0 = [X\ Y\ Z]^T$

↳ Location of camera in world frame

$\theta = [\alpha\ \beta\ \gamma]^T$

↳ Where your camera is looking at

$X_0 + \theta \rightarrow$ 6 vectors → Give the 6 DOF

↓

Extrinsics

↓

Help us perform camera localization

Intrinsics:

- $c \rightarrow$ Camera constant
  ⇓
  Dist. of image plane to projection center
- $m \rightarrow$ Scale difference b/w x-y axes
- $f_x, f_y \rightarrow$ Focal lengths in x, y directions
- $x_H, y_H \rightarrow$ Principal point
  - $x_H \rightarrow u_0$, $y_H \rightarrow v_0$
  - Principal point ↓ Pt. thru which optical axis of camera passes
- $s \rightarrow$ Shear parameter (For analog camera)
- These 4 or 5 params describe camera intrinsics

Direct Linear Transform (DLT):

- 11-DOF transform → 6 from Extr. + 5 from Intr.
- Lens distortion assumed to be absent
  ↳ Lens is perfect
- 6 or more control points
- We have barrel, cushion distortions in lens distortion
  ↳ Add non-linear params to camera model
- $x = PX \rightarrow$ 3D world coordinates
  - x ↓ 2D coords
  - P ↳ Projection matrix takes from 3D to 2D
  - All coords in homogeneous coords
  - P has all the 11 params
  - Intrinsics from Cam Calib
  - Extrinsics from Cam localization
- Projection cam model → All rays intersect at projection center

Coordinate systems:

- **World:** World frame in 3D
- **Camera:** 3D frame with origin at cam projection center
- **Image plane:** Plane onto which 3D coords are projected to
- **Sensor:** Has pixel locations
- Projection center → (X_0, R)
  - X_0 ↓ Location of proj. center in world frame
  - R ↳ Direction in which cam is looking at

$$P_{3\times4} = K_{3\times3} \underbrace{\begin{bmatrix} R & -RX_0 \\ \cancel{\quad} & \end{bmatrix}_{3\times4}}$$

World to img ~~frame~~ frame

Depth not lost yet

Can go back to world frame at this instant

K takes from 3D img frame to 2D cam frame

Depth is lost → $\boxed{\text{Direct inversion of P not possible}}$

$$\boxed{P_{3\times4} = K_{3\times3}\begin{bmatrix} R & -RX_0 \end{bmatrix}_{3\times4}}$$

$$\boxed{P_{3\times4} = K_{3\times3}\, R_{3\times3} \underbrace{\begin{bmatrix} I / -X_0 \end{bmatrix}_{3\times4}}}$$

The translated point is then rotated

After the transformation to img frame is complete, we project to pixel coords

Translates world coords such that new origin is proj. center

Translation matrix from homo coords notes

line on which pixel lies can be found out tho

This helps recovery of 3D coord in multi view geo as they are found from intersection of these lines for the same point across views

Direct Linear Transform (DLT):

- To estimate cam localization as well as calib params
- Need at least ~~*~~ 6 pts in world frame and a single picture of these points
- Estimates R, X_0, K

$$\underbrace{R}_{3 \text{ params}}, \quad \underbrace{X_0}_{3 \text{ params}}, \quad \underbrace{K}_{5 \text{ params}}$$

- $\underset{3\times1}{x} = \underset{3\times4}{P} \, X_{4\times1}$

$P = KR[I \mid -X_0]$

$P_{3\times4} \rightarrow 12$ values

- $a_x, a_y \rightarrow$ Coeff vectors for x and y coords
- $\boxed{\begin{array}{l} a_x^T p = 0 \\ a_y^T p = 0 \end{array}}$ $\rightarrow$ Projection vector with 12 values

- 2 coeff vector per point

⇓

$\boxed{\text{2 equations per point}}$

&

12 params in p vector

⇓

At least 6 points needed

- 12 equations into matrix $M_{12\times12}$

$M_{12\times12} \, p_{12\times1} = 0$

$\hookrightarrow \boxed{\text{SVD}} \rightarrow p$ which minimises

⇓ $\boxed{\text{QR decomp.}}$

R, X_0, K