11. Camera Calibration Procedure

09 February 2024 12:11

Source: First Principles of Computer Vision, Prof. Shree Nayyar

From <https://fpcv.cs.columbia.edu/>

Mathematics :-

Linear Model Approximation :-

Linear Model

$$\underset{\substack{m\times 1\\(\text{known})}}{y} = \underset{\substack{m\times n\\(\text{matrix})}}{A}\ \underset{\substack{n\times 1\\(\text{unknown})}}{x}$$

m = # of equations

n = # of unknowns

If $m = n$ and A is invertible.

$$\boxed{x = A^{-1} y}$$

If $m > n$

# of equations > # of unknowns

$\Rightarrow$ overdetermined system.

Using least squares solution :-

$$\boxed{x = (A^T A)^{-1} A^T y}$$

Noise in the system.

$\Rightarrow$ No exact solutions.

Maximum - Likelihood Solution :-

$$y = Ax + \underset{\text{noise}}{n}$$

$$\Rightarrow y \neq Ax$$

$$y - Ax = e$$

error vector (Approximation)

To find best x,

minimize approximation error

$$\min \|e\|$$
$$= \min \|y - Ax\|$$
$$= \min \|y - Ax\|_2^2$$

(Least Squares Problem)

$$\min \|y - Ax\|^2 = (y - Ax)^T \cdot (y - Ax)$$
$$= (y - x^T A^T) \cdot (y - Ax)$$
$$= y^T y - \underline{x^T A^T y} - \underline{y^T A x} + x^T A^T A x$$

$$f(x) = y^T y - 2 x^T A y + x^T A^T A x$$

Taking gradient,

$$\nabla_x f(x) = 0 - 2 A^T y + 2 A^T A x$$

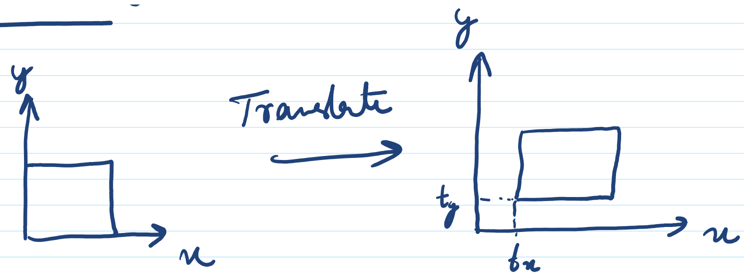
$$A^T A x = A^T y$$

$$\boxed{x = (A^T A)^{-1} A^T y}$$

Least squares solution assuming $A^T A$ is invertible.

Homogeneous Coordinates :-

y



$$x_2 = x_1 + t_x$$
$$y_2 = y_1 + t_y$$

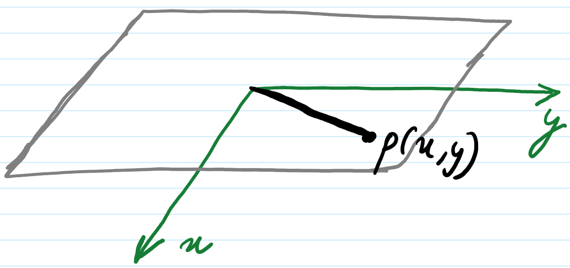
Can we use 2x2 matrix here ?

$$\begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \tilde{x}_2 \\ \tilde{y}_2 \\ \tilde{z}_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & t_x \\ 0 & 1 & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

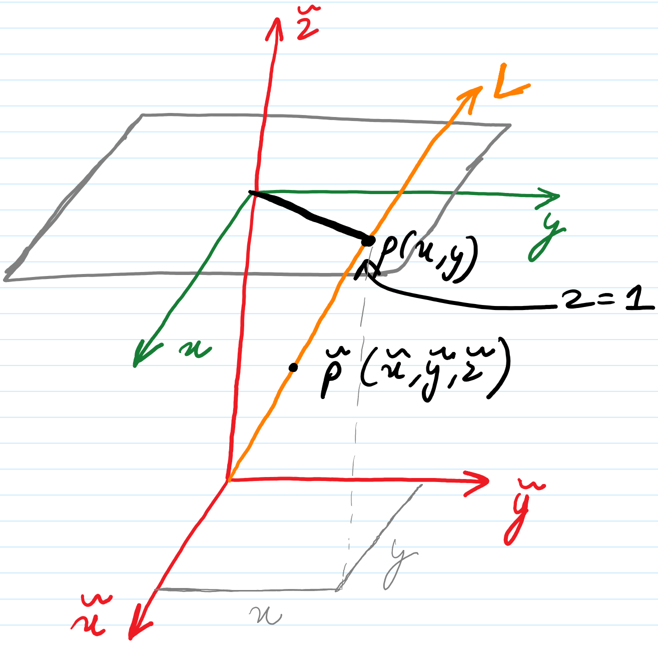
Intuitively, we are representing the same points in one dimension up.

It is a homogeneous representation of a 2D point $p = (x, y)$ as a 3D point $\tilde{p} = (\tilde{x}, \tilde{y}, \tilde{z})$. The third coordinate $\tilde{z}$ $(\neq 0)$ is such that

$$x = \frac{\tilde{x}}{\tilde{z}}, \quad y = \frac{\tilde{y}}{\tilde{z}} \longleftarrow \text{fictitious}$$



$\wedge \tilde{z}$

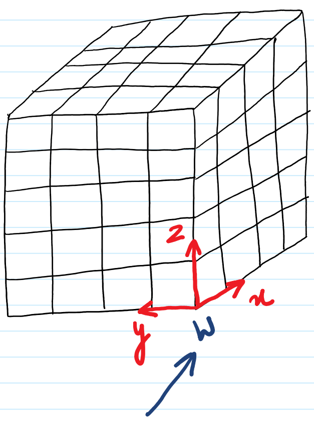


$$p = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \equiv \begin{bmatrix} \tilde{z}x \\ \tilde{z}y \\ \tilde{z} \end{bmatrix} \equiv \begin{bmatrix} \tilde{x} \\ \tilde{y} \\ \tilde{z} \end{bmatrix} = \tilde{p}$$

Every point on L (>0) represents homogeneous coordinates of $p(x,y)$

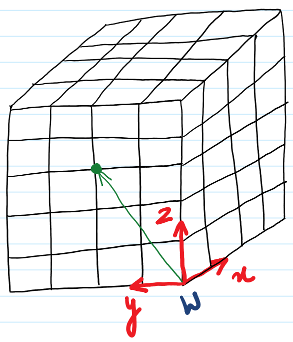
⇒ Camera Calibration Procedure :-

① Capture an image of an object with known geometry.

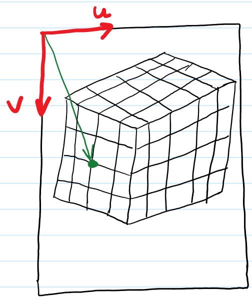


An object whose corner is the world coordinate frame.

② Identify correspondences between 3D scene points and image points.



Object of known geometry



Captured Image

$$X = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \\ 4 \end{bmatrix} \qquad u = \begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} 56 \\ 115 \end{bmatrix}$$

(inches) (pixels)

③ For each corresponding point i in the scene and image:

$$\underbrace{\begin{bmatrix} u^{(i)} \\ v^{(i)} \\ 1 \end{bmatrix}}_{\text{known}} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\text{unknown}} \underbrace{\begin{bmatrix} x_w^{(i)} \\ y_w^{(i)} \\ z_w^{(i)} \\ 1 \end{bmatrix}}_{\text{known}}$$

Expanding,

$$\ldots^{(i)} \quad x_w^{(i)} + p_{12} y_w^{(i)} + p_{13} z_w^{(i)} + p_{14}$$

expanding,

$$u^{(i)} = \frac{p_{11}\, x_w^{(i)} + p_{12}\, y_w^{(i)} + p_{13}\, z_w^{(i)} + p_{14}}{p_{31}\, x_w^{(i)} + p_{32}\, y_w^{(i)} + p_{33}\, z_w^{(i)} + p_{34}}$$

$$v^{(i)} = \frac{p_{21}\, x_w^{(i)} + p_{22}\, y_w^{(i)} + p_{23}\, z_w^{(i)} + p_{24}}{p_{31}\, x_w^{(i)} + p_{32}\, y_w^{(i)} + p_{33}\, z_w^{(i)} + p_{34}}$$

④ Rearranging

$$\underbrace{\begin{bmatrix} \quad & \quad \\ \quad & \quad \end{bmatrix}_{12\times n}}_{\text{known}} \underbrace{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{13} \\ \vdots \\ p_{34} \end{bmatrix}_{12\times 1}}_{\text{unknown}} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{12\times 1}$$

⑤ Solve for $Ap = 0$

Decomposition of Project matrix P
into intrinsic and extrinsic parameters :-

We know that :-

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{M_{int}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{M_{ext}}$$

We know,

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \end{bmatrix} = \ldots$$

We know,

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix} = \begin{bmatrix} f_x & 0 & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix} \boxed{\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}}$$

$$= K R$$

upper triangular

orthogonal

Using "QR factorization"

Also,

$$P = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} p_{14} \\ p_{24} \\ p_{34} \end{bmatrix} = \begin{bmatrix} f_x & 0 & O_x \\ 0 & f_y & O_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} = K t$$

$$t = K^{-1} \begin{bmatrix} p_{14} \\ p_{24} \\ p_{34} \end{bmatrix}$$

Forward Projection : 3D $\rightarrow$ 2D

Backward Projection : 2D $\rightarrow$ 3D

$\hookrightarrow$ Given a calibrated camera, can we find the 3D scene point from a single 2D image?