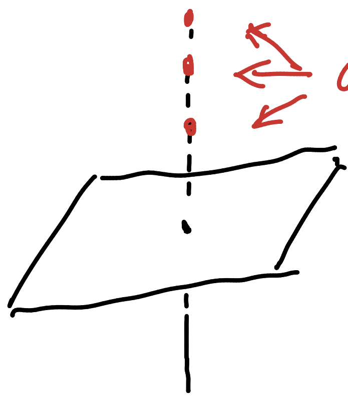
lect 13 Recover 3D-structure from image

Actually, it's ambiguous, because



→ One method

Multi-view.

Normalized (Camera) Coordinate System

O is the pin-hole x-y plane parallel to film. z-axis perp to film.

1. change arbitary coordinate system to image coordinate system. for pin-hole camera model

$(X, Y, Z) \mapsto (fX/Z, fY/Z)$

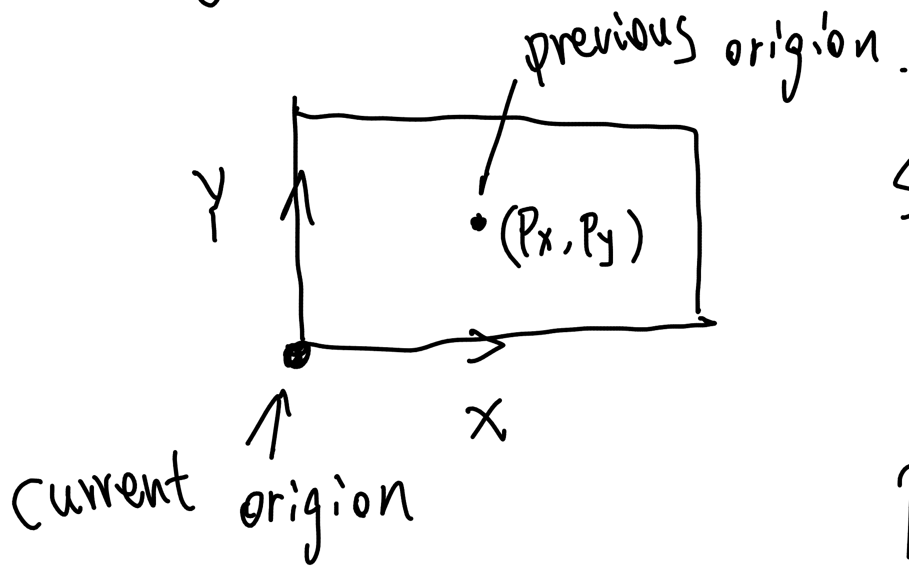
Homogenious coordinate

$$\underset{3D}{\begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}} \overset{map}{\mapsto} \underset{2D}{\begin{pmatrix} fX \\ fY \\ Z \end{pmatrix}} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix} \Rightarrow \text{so } X' = P \cdot X$$

Principle point P.

the axis perp to film is call principle axis. The projection of O point is principle point.

Usually we have the following system



So, $(X, Y, Z) \mapsto (fX/Z + \underbrace{P_x}_{\text{offset}}, fY/Z + \underbrace{P_y}_{\text{offset}})$

Now
$$\begin{pmatrix}X\\Y\\Z\\1\end{pmatrix} \mapsto \begin{pmatrix} fX+P_xZ\\ fY+P_yZ\\ Z\end{pmatrix} = \overbrace{\begin{pmatrix} f & & \boxed{P_x} & 0\\ & f & \boxed{P_y} & 0\\ & & 1 & 0\end{pmatrix}}^{\text{Projection matrix}} \begin{pmatrix}X\\Y\\Z\\1\end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} f & & P_x\\ & f & P_y\\ & & 1\end{pmatrix}}_{K} \begin{pmatrix} 1 & & & 0\\ & 1 & & 0\\ & & 1 & 0\end{pmatrix} \begin{pmatrix}X\\Y\\Z\\1\end{pmatrix}$$

K : Calibration matrix,

contain all info ~~of~~ of camera

Pixel Coordinate

Pixel size $= \frac{1}{m_x} \times \frac{1}{m_y}$

$\uparrow$ # of pixels per meter, or called density

$$K = \begin{bmatrix} m_x & & \\ & m_y & \\ & & 1 \end{bmatrix} \begin{bmatrix} f & & p_x \\ & f & p_y \\ & & 1 \end{bmatrix} = \begin{bmatrix} f m_x & & p_x m_x \\ & f m_y & p_y m_y \\ & & 1 \end{bmatrix} = \begin{bmatrix} \alpha_x & & \beta_x \\ & \alpha_y & \beta_y \\ & & 1 \end{bmatrix}$$

# of pixel/meter

K turns to connect tightly with pixel (may not be integer)

Rotation & Translation.

World coordinate system may need rotation & translation to turn to camera coordinate system

$$\tilde{X}' = R(\tilde{X} - \tilde{C})$$

Non-homogenious system marked with "$\sim$"

rotation

image point in world system

camera center in world system

$$X' = \begin{pmatrix} \tilde{X}' \\ 1 \end{pmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \tilde{X} \\ 1 \end{bmatrix} = \begin{bmatrix} R & -R\tilde{C} \\ 0 & 1 \end{bmatrix} X$$

Now, our final projection is

$x = K[I|0]X_{cam}$ Camera system

$= K[R| \underbrace{-R\tilde{C}}_{t_{3\times1}}]X$ world system

$= K[R|t]X$

$= PX$

$\boxed{\text{fact: when } X = \begin{bmatrix}\tilde{C}\\1\end{bmatrix},\ PX = 0, \text{ so } \begin{bmatrix}\tilde{C}\\1\end{bmatrix} \text{ is the kernel of } P}$

This system doesn't consider

* skew (Non-rectangular pixel)
* Radial (line → curve)

So, once we find $\underline{P} \in \mathbb{R}^{3\times4}$, or $K[R|t]$, we are done.

Method 1. Linear Method

$$\lambda\begin{bmatrix}x\\y\\1\end{bmatrix} = P\begin{bmatrix}x\\y\\z\\1\end{bmatrix} \Leftrightarrow \lambda_i x_i = PX_i \Leftrightarrow x_i \times PX = 0$$

$$\begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \times \begin{bmatrix} P_1^T \underline{X}_i \\ P_2^T \underline{X}_i \\ P_3^T \underline{X}_i \end{bmatrix} = 0$$

[Similar topic see lect 9]

$$\begin{bmatrix} 0 & -\underline{X}^T & -y\underline{X}^T \\ \underline{X}^T & 0 & -x\underline{X}^T \\ -y\underline{X}^T & x\underline{X}^T & 0 \end{bmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix}_{12\times 1} = 0 \qquad P \sim 11 \text{ DoF}$$

to avoid lazy dummy solution, add $\|P\|_2 = 1$

try not to choose points lying on the same plane, say.

$(0, \text{somepoints}, 0)^T$, because we need 11-DoF

Pros: EASY!

Cons: Doesn't directly tell you what is f, m_x, R, $\tilde{c}$, etc

Doesn't remove radial distortion.

If we know some parameters, no use

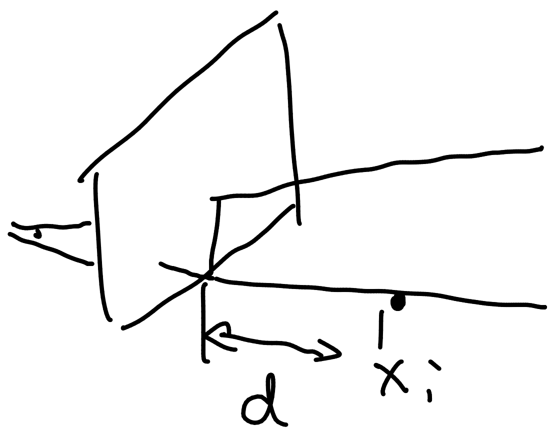
Method 2. Non-linear Method

Define squared distance,

Use Newton's method to solve. (not mentioned

What if we don't have ground truth for camera calibration

→ We can make use of cues from: vanish point



$$\underline{X}_t = \begin{bmatrix} x_0 + t\,d_x \\ y_0 + t\,d_y \\ z_0 + t\,d_z \\ 1 \end{bmatrix} \xrightarrow{t \to \infty} \begin{bmatrix} dx \\ dy \\ dz \\ 0 \end{bmatrix}$$

$[dx, dy, dz]$ is the line direction

The vanish point $V = P X_\infty$

if $\begin{cases} V \text{ is a finite point} & V = \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} \\ V \text{ is a infinite point} & V = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \end{cases}$

Now if we choose a point on $(1, 0 0 0)^T$, it is the infinite point on X-axis.

$P = [P_1, P_2, P_3, P_4]_{3\times 4}$ $\quad P \cdot [1\ 0\ 0\ 0]^T = P_1$ → vanish point in x direction

P_2, P_3 are y, z's vanish point

P_4 is origion of world coordinate system, because its $(0,0,0)$ point in real world. ($P_4 = P[0\ 0\ 0\ 1]^T$)

But, we need to know the scaling of each column, $P_1 \cdots P_4$,

let $e_1 = \begin{bmatrix}1\\0\\0\end{bmatrix}$ $e_2 = \begin{bmatrix}0\\1\\0\end{bmatrix}$ $e_3 = \begin{bmatrix}0\\0\\1\end{bmatrix}$ they are the expected basis for 3D world

$$\lambda_i v_i = K[R|t]\begin{bmatrix}e_i\\0\end{bmatrix} = KRe_i$$

$$\Rightarrow e_i = \lambda_i R^T K^{-1} v_i \quad \text{and} \quad e_i^T e_j = 0$$

So, $$e_i^T e_j = 0 \iff \underbrace{\lambda_i}_{\text{scalar}} v_i^T K^{-T} \underbrace{RR^T}_{I} K^{-1} v_j \underbrace{\lambda_j}_{\text{scalar}} = 0$$

$$\boxed{v_i^T K^{-T} K^{-1} v_j = 0}$$

$\rightarrow$ for $i = 1, 2, 3$. we have 3- equation set.

$\Rightarrow$ solve out $\underline{\underline{K}}$

But

if we have: (It doesn't matter for other cases)

1 finite vanishing point
2 - infinite v-p. $\Big] \Rightarrow$ we could not recover focus length

---

$\lambda_i V_i = KRe_i \Rightarrow Re_i = \lambda K^{-1} V_i \Rightarrow$ we get a column of R

$\uparrow$ Known

Advantage:

- No need for calibration chart.
- Can be automatically done

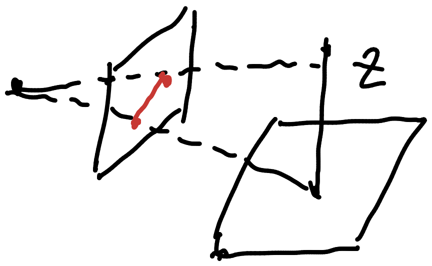
Disadvantage.

- We should find vanishing points. limit of application
- Inaccurate to compute vanishing points.
- When having two infinite vanishing points. done...

Measure Height.

Recall, we could use vanishing line to measure height.

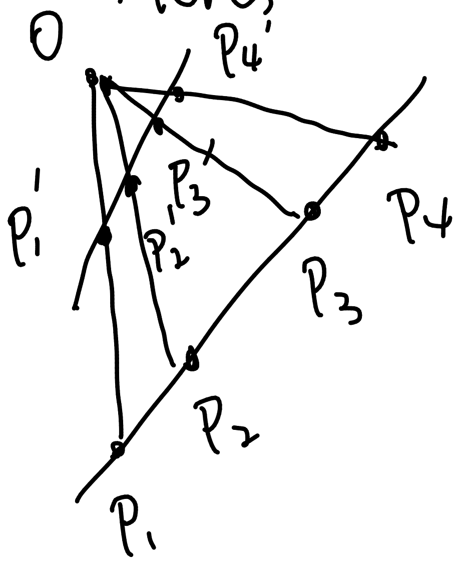
How to measure height w/o ruler ?



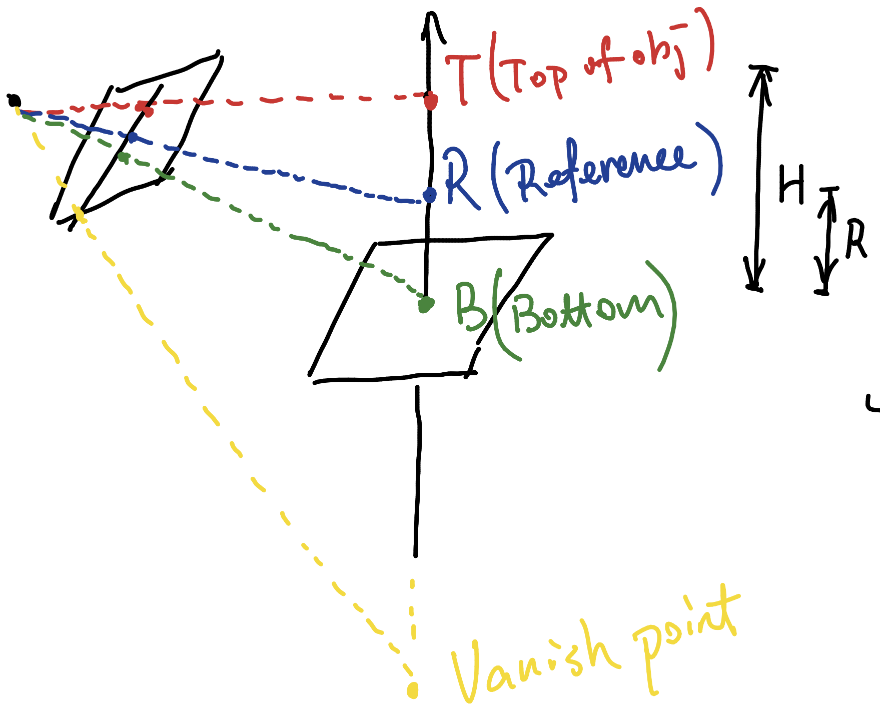
Cross-ratio

$\boxed{\text{Projective Invariant}}$: Some quantity that won't change after projection.

Here, we use the "ratio" of length. (2D or 3D, applicable)



$$\frac{\|P_3 - P_1\| \|P_4 - P_2\|}{\|P_3 - P_2\| \|P_4 - P_1\|} \text{ is constant}$$



$$\frac{\|T-B\|\,\|R-\infty\|}{\|R-B\|\,\|T-\infty\|} = \frac{\|T-B\|}{\|R-B\|}$$ Real World

$$\parallel$$

$$\frac{\|t-b\|\,\|R-V_z\|}{\|r-b\|\,\|T-V_z\|} = \frac{H}{R}$$ Image World

Vanishing point

2D-lines in homogenious coordinate system

line equation $ax + by + c = 0$

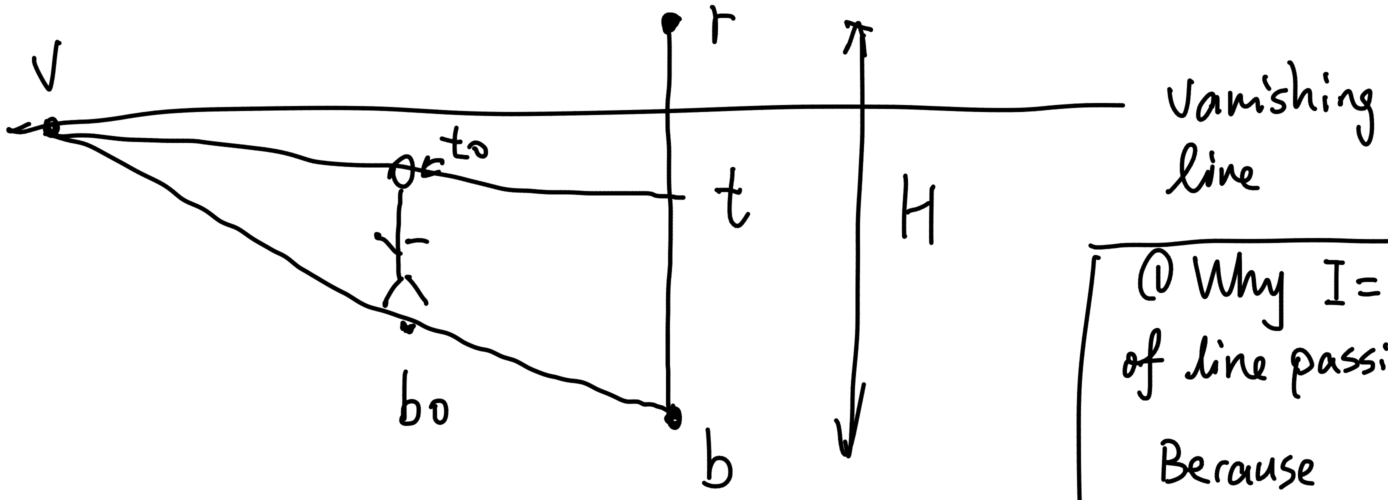
$$\Rightarrow \begin{bmatrix} a \\ b \\ c \end{bmatrix}^T \begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = 0$$

$\begin{bmatrix} a \\ b \\ c \end{bmatrix}$ is the normal direction of the line

① If we have two points $\begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}, \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$

the coefficients $l = \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix} \times \begin{bmatrix} x_2 \\ y_2 \\ 1 \end{bmatrix}$ is a line passing them

② two lines: l_1, l_2, the intersection $x = l_1 \times l_2$



$$V = (\underbrace{b \times b_0}_{\text{line cross } b, b_0}) \times (\underbrace{V_x \times V_Y}_{\text{Vanishing line}})$$

line cross b, b0

Vanishing line

$$t = (\underbrace{v \times t_0}_{\text{line pass } v \,\&\, t_0}) \times (\underbrace{r \times b}_{\text{line pass } r, b})$$

line pass v & to

line pass r, b

① Why $I = X_1 \times X_2$ is the coefficient of line passing point X_1, X_2?

Because

$I = X_1 \times X_2 \iff I^T \cdot X_1 = I^T X_2 = 0$

so, I is the normal vector of the line.

② Why $X = I_1 \times I_2$ is the intersection of two lines?

Because.

$X = I_1 \times I_2 \iff x^T I_1 = x^T I_2 = 0$

so, x satisfy both line function

If we would like to measure a line on the plane?

1. Unwarp the image: Use Homography Transformation to turn the plane parallel to film, then measure. (let the line // film)

We need four points on the plane to compute the H of the Homography transformation (rectify method).