Camera Calibration :

→ How to recover 3D structure of the scene from its images?

→ To go from images and full metric reconstruction we need two thing→

1)→ first is position and orientation of camera with respect to the world and coordinate frame. → These are external parameters

2)→ $\boxed{\text{interal parameter}}$ such as focal length

How camera maps their perspective projection on its image plane

→

Method to find camera's internal and external parameters

1)→ Linear Camera Model. $\boxed{\text{3D से px in img}}$

turns out to be a single matrix called projection matrix.

हमे सिर्फ known geometry के object की zarurat है जिससे camera calibrate हो जायेगा।

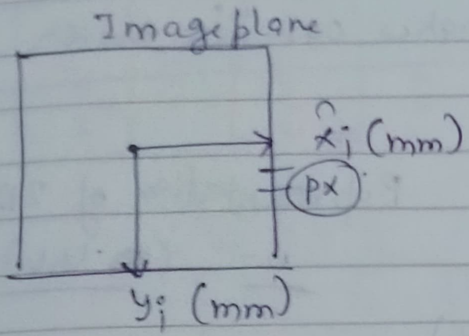
→ To determine to projection matrix

intrinsic and Extrinsic matrix pata लगाना है।

2025/03/02 21:38

perspective projection equations

$$\boxed{\frac{x_i}{f} = \frac{x_c}{z_c}} \qquad \boxed{\frac{y_i}{f} = \frac{y_c}{z_c}}$$



$x_i \rightarrow$, $y_i \rightarrow$ projection of point P onto image plane

px could be rectangular, it depends on px density.

$\boxed{m_x} \rightarrow$ px density in x

$\boxed{m_y} \rightarrow$ px density in y dir^n

$$\boxed{\begin{aligned} u &= m_x x_i \\ &= m_x \frac{x_c f}{z_c} \end{aligned}} \qquad \boxed{v = m_y y_i = m_y \left(\frac{y_c}{z_c}\right) f}$$

we don't know the point of image piercing the plane, usually it is taken at the corner of the plane

that's modify our equation.

$$\boxed{u = m_x f \frac{x_c}{z_c} + o_x} \qquad v = m_y f \left(\frac{y_c}{z_c}\right) + o_y$$

combining

$$\boxed{u = f_x \left(\frac{x_c}{z_c}\right) + o_x} \qquad v = f_y \left(\frac{y_c}{z_c}\right) + o_y$$

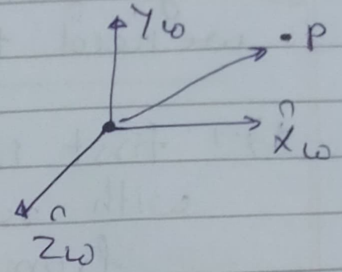
f_x → effective focal length

2025/03/02 21:38

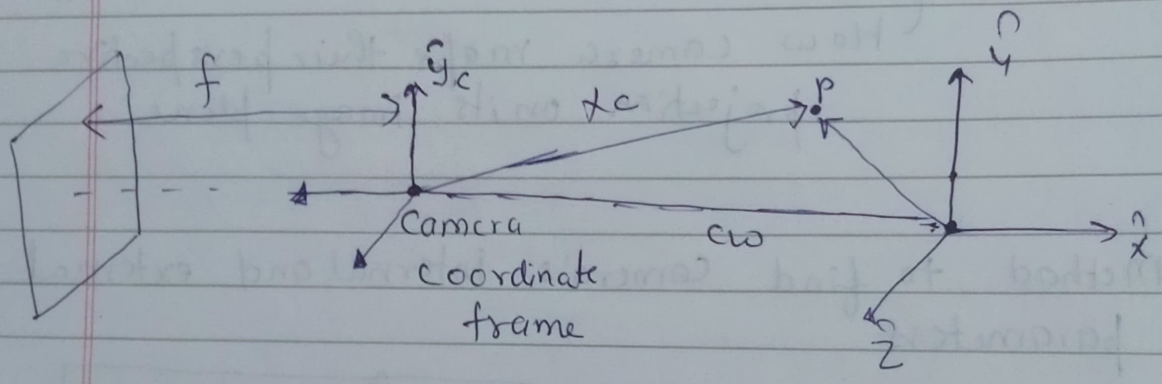
$\boxed{\text{Simple Stereo}}$ ← How to recover 3D structure from 2 calibrated camera.

→ $\boxed{\text{Forward Imaging Model}}$

it takes us from 3D to 2D.



world coordinate frame



Z_c → z axis of C.CF is aligned with ~~cam~~ optical Axis

अगर हमे CCF पता wrt world C.F तो हम उसकी mapping find कर सकते है। projection create कर सकते है

$$x_i = \begin{bmatrix} x_i \\ y_i \end{bmatrix} \qquad \underline{x_c} = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} \qquad x_w = \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix}$$

if you have x_c → use करे $\boxed{\text{perspective projection}}$

perspective projection.

$$u = f_x\left(\frac{x_c}{z_c}\right) + o_x \quad \text{—(1)} \qquad v = f_y\left(\frac{y_c}{z_c}\right) + o_y$$

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} \equiv \begin{bmatrix} z_c u \\ z_c v \\ z_c \end{bmatrix} \quad (2)$$

where

$$z_c u = \tilde{u}$$

$$\boxed{z_c = \frac{\tilde{u}}{u}}$$

$$= \quad \vec{w}\, u = \tilde{u}$$

$$\boxed{u = \frac{\tilde{u}}{\tilde{w}}} \qquad \boxed{v = \frac{\tilde{v}}{\tilde{w}}}$$

on multiplying 1 by z_c

$$z_c \cdot u = f_x \cdot x_c + o_x z_c = \begin{bmatrix} f_x x_c + o_x z_c \\ f_y y_c + z_c o_y \\ z_c \end{bmatrix}$$

simply karo

$$\underbrace{\begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\text{it include all the internal parameter}} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

is ~~sin~~ ko simplify करो।
लिखना।

it include all the internal parameter

this is linear model for perspective projection

2025/03/02 21:39

$(f_x\ f_y\ O_x\ O_y) \rightarrow$ Intrisinic parameters of the camera

They represent the camera's internal geometry.

$\boxed{\text{Go to linear model}}$ $\boxed{\text{it makes our life easier.}}$

$\boxed{\tilde{w} \neq 0}$

$\boxed{\text{Homogenous Coordinates}}$ Representation of 2D point
$u = (u, v)$

is 3D point $\tilde{u} = (\tilde{u}, \tilde{v}, \tilde{w})$. w is fictitious

$\boxed{u = \frac{\tilde{u}}{\tilde{w}}}$ $\boxed{v = \frac{\tilde{v}}{\tilde{w}}}$ $u = (uv, v, 1)$

used for scaling and normalization

$\tilde{u}' = u', v', w'$

$\boxed{u = \frac{u'}{w'}}$

$\boxed{v = \frac{u'}{w'}}$

$$u = \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \tilde{w} u \\ \tilde{w} v \\ \tilde{w} \end{bmatrix} = \begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} = \tilde{u}$$

Same you can do with 3D → 4D point

2025/03/02 21:38

Rational matrix is orthonormal → $\boxed{dot(u,v)=0}$

$\boxed{u^T v = 0}$ incase of matrix

$\boxed{u \cdot v = 0}$

$\boxed{\begin{matrix} u^T u = 1 \\ v^T v = 1 \end{matrix}}$ → unit length

eg)- x axix, y axix, z axis

In orthonormal matrix → square matrix whose row/col vectors are orthonormal.

$\boxed{R^{-1} = R^T}$ ✓ $\boxed{R^T R = R R^T = I}$ condition of normalatity

→ $\boxed{x_c = R(x_w - c_w)}$ find x_c in world coordinate frame

C-CF, R.m, wrt WCF

$x_c = (x_w - c_w) R$

we have mapped to wcf to ccf

$\boxed{t = -R c_w}$ translation matrix

$= x_c = R x_w - R c_w = R x_w + t$

$$x_c = \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$$

यहां एक Doubt है

wcf → world coordinate frame

Intrinsic Matrix

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{u} \\ \bar{v} \\ \bar{w} \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

$$K = \begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix} \rightarrow \text{calibration Matrix}$$

it is upper Right Triangular matrix → import

$M_{int} = [K|0]$ → concatenation of $K|0$

$$\begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

mapping the world coordinates to camera coordinate

$\boxed{3D \text{ to } 3D}$ → The position and orientation to camera coordinate frame to world coordinate frame

C_w → position

R → orientation

3x3 matrix

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

→ Dir^n of $\hat{x}_c$ in WCF

→ Dir^n of $\hat{y}_c$ in WCF

→ Dir^n of $\hat{z}_c$ in WCF

2025/03/02 21:39

$$x_c = \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

isko solve karke dekhna ho

$$M_{ext} = \begin{bmatrix} R_{3\times3} & t \\ 0_{1\times3} & 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\boxed{\tilde{x}_c = M_{ext} \tilde{x}_w}$$

projection matrix P

$$\begin{bmatrix} \tilde{u} \\ \tilde{v} \\ \tilde{w} \end{bmatrix} = \begin{bmatrix} f_x & 0 & o_x & 0 \\ 0 & f_y & o_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix}$$

$$\boxed{\tilde{u} = M_{int} \tilde{x}_c}$$

$$\begin{bmatrix} x_c \\ y_c \\ z_c \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

$$\boxed{\tilde{x}_c = M_{ext} \tilde{x}_w}$$

2025/03/02 21:39

set of corresponding point ले जायेंगे | → map करना होगा |

$\boxed{\text{Step 3}} \rightarrow \underbrace{\begin{bmatrix} u_i \\ v_i \\ 1 \end{bmatrix}_{3\times1}}_{\boxed{\text{known}}} = \underbrace{\begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix}}_{\text{unknown } 3\times4} \underbrace{\begin{bmatrix} x_w^i \\ y_w^i \\ z_w^i \\ 1 \end{bmatrix}_{4\times1}}_{\text{known}}$

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

$$v_i = \frac{p_{21} x_w^{(i)} + p_{22} y_w^i + p_{23} z_w^i + p_{24}}{p_{31} x_w^i + p_{32} y_w^1 + p_{33} z_w^i + p_{34}}$$

$$\begin{bmatrix} A \end{bmatrix} \begin{bmatrix} p_{11} \\ p_{12} \\ \vdots \\ p_{34} \end{bmatrix} = 0$$

$\boxed{AP=0}$ equation का screenshot ले लिया है |

~~Sa~~ Scaling का ~~kum~~ kch impact नहीं पड़ता है | $\boxed{k=1}$ या $\underline{\|p\|^2=1}$ भी हो सकता है |

Combining →

$$\boxed{\tilde{u} = M_{int} M_{ext}\ \tilde{x}_w = P\ \tilde{x}_w}$$

$$\boxed{P = M_{int} M_{ext}}$$

image

points in 3D

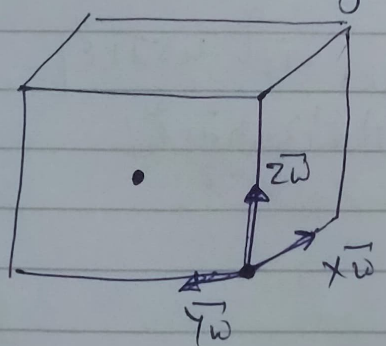
$$\begin{bmatrix} u \\ v \\ w \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \end{bmatrix} \begin{bmatrix} x_w \\ y_w \\ z_w \\ 1 \end{bmatrix}$$

We only need to find projection matrix to calibrate camera

Lecture 3

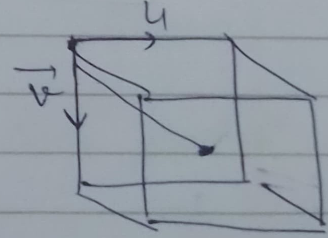
estimation of Projection matrix.

1) → capture an image of an object with known geometry.



place WCF in one of the corner of the cube

image बना लो



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

WCF inches px

$$\begin{bmatrix} P_{14} \\ P_{24} \\ P_{34} \end{bmatrix} = \underbrace{\begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K} \underbrace{\begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}_{t} .$$

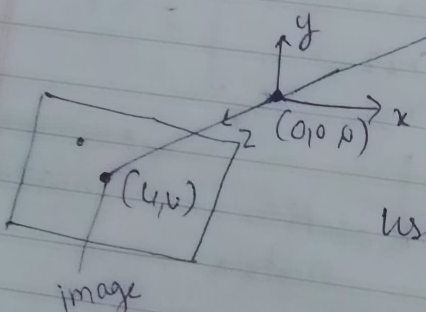
$$\boxed{t = K^{-1} \begin{bmatrix} P_{14} \\ P_{24} \\ P_{34} \end{bmatrix}} \rightarrow \text{translation matrix}$$

- lenses have Distortion
  - Radial
  - Tangential

Pinhole में Distortion नहीं होता है।

it makes the model non-linear

Simple Stereo → Recover 3D structure with 2D image



3D point it must lie on any outgoing ray. we know the equation of this ray

using perspective projection we can find out

$$u = f_x \frac{x_c}{z_c} + o_x \qquad v = f_y \frac{y_c}{z_c} + o_y$$

2D to 3D

by rearranging this

$$\boxed{\frac{(u - o_x)\, z_c}{f_x} = x}$$

$$\boxed{z > 0}$$

y भी निकालो

Intrinsic and Extrinsic matrix

projection matrix can be decomposed further into intrinsic / Extrinsic matrix

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

Consider only this

$$\begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} f_x & 0 & o_x \\ 0 & f_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\text{calibration matrix}} \underbrace{\begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}}_{\text{Rotation matrix}} = \underline{\underline{KR}}$$

orthonormal matrix

it is possible to find K and R from their product using "QR factorization".

Find translational now.

$$P = \begin{bmatrix} - & - & - & p_{14} \\ - & - & - & p_{24} \\ - & - & - & p_{34} \end{bmatrix} =$$

transform it

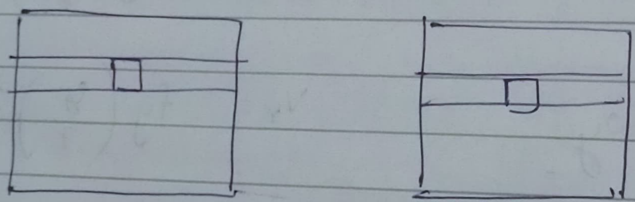
see next page

2025/03/02 21:39

- $\boxed{\text{Disparity} \propto \text{baseline}}$ as per the formula
- larger the baseline, more precisely ~~at~~, disparity is calculated.
- Stereo Matching
  - Find the Disparity between left and right stereo pair.

there is no disparity in Y dir^n in case of Horizontal disparity

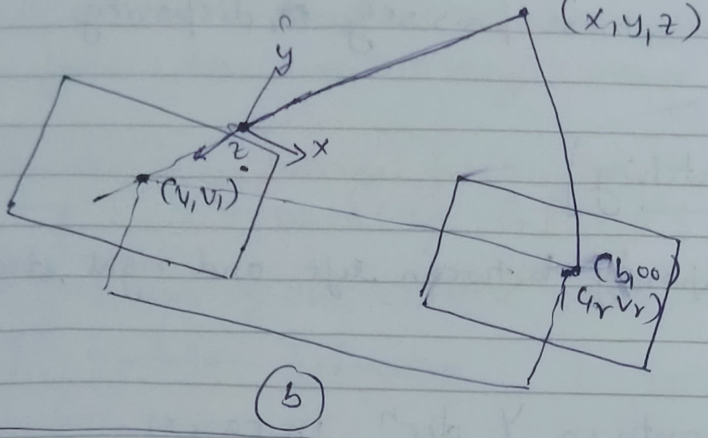
$$\boxed{v_l = v_r = f_y\left(\frac{y}{z}\right) + o_y}$$ — Each corresponding.



By this we known that this portion must lie on the same horizontal ~~line~~. | window

Issues
- texture होना चाहिए]
- Should ~~be~~ not be repetative
- window size होना चाहिए]

We need more info to construct 3D ~~image~~ structure out of image



This system is known as simple stereo binocular vision

We have 4 eqn →

$u_l = f_x \frac{x}{z} + o_x$

$u_R = f_x \left(\frac{x-b}{z}\right) + o_x$

$v_l = f_y \, y/z + o_y$

$v_r = f_y \left(\frac{y}{z}\right) + o_y$

Solving for (x, y, z) →

$x = \frac{b(u_l - o_x)}{u_l - u_r}$

$y = \frac{b f_x (v_l - o_y)}{f_y (u_l - u_r)}$

imp

$$\boxed{z = \frac{b f_x}{u_l - u_r}}$$

depth →

$u_l - u_r$ → Difference in the u coordinate. (Disparity)

disparity is inversly proportinal to z

Uncalibrated ~~camera~~ Stereo →

Method to estimate 3D structure of a static scene from two arbitrary views.

assume → internal parameters are known to us
↳ metatag में होता है।

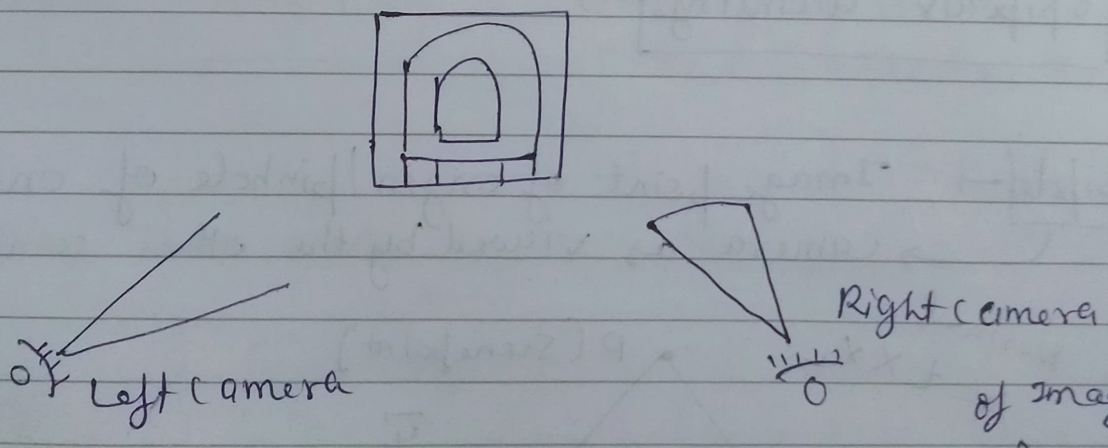
1) problem of uncalibrated stereo.

1) Epipolar Geometry | Estimating fundamental matrix

→ finding dense correspondences

→ Computing Depth

1) problem of uncalibrated Stereo.



Instrincs are known to us → meta tag of image में से in होता है
$(f_x, f_x\ o_x, o_y)$

Extrincs (~~rep~~ relative position | orientation of cameras) unknown.

$$[x_l \; y_l \; z_l] \begin{bmatrix} t_y z_l - t_z y_l \\ t_z x_l - t_x z_l \\ t_x y_l - t_y x_l \end{bmatrix} = 0$$

Cross-product defination

↓

matrix मे भी convert कर सकते है

$$\underbrace{[x_l \; y_l \; z_l] \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_x & t_x & 0 \end{bmatrix}}_{T_x \to \text{Translation matrix}} \begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = 0 \quad \text{--- (1)}$$

$t_{3\times1} \to$ position of Right camera in Left Camera's frame

$R_{3\times3}$ " Orientation "

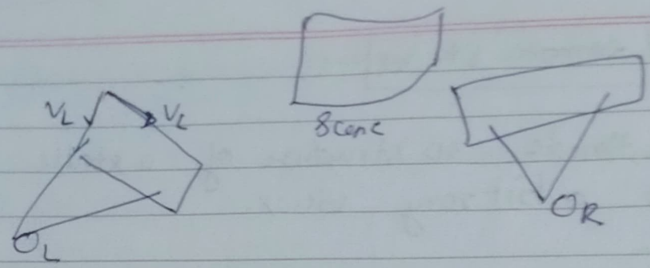
$$\boxed{x_l = R x_r + t}$$

$$\begin{bmatrix} x_l \\ y_l \\ z_l \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \quad \text{--- (2)}$$

Substituting 2 eq into 1.

$$[x_l \; y_l \; z_l] \Bigg(\overbrace{\begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}}^{\text{essential m.}} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix}$$

$$+ \underbrace{\begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}}_{\boxed{t \times t = 0}} = 0$$

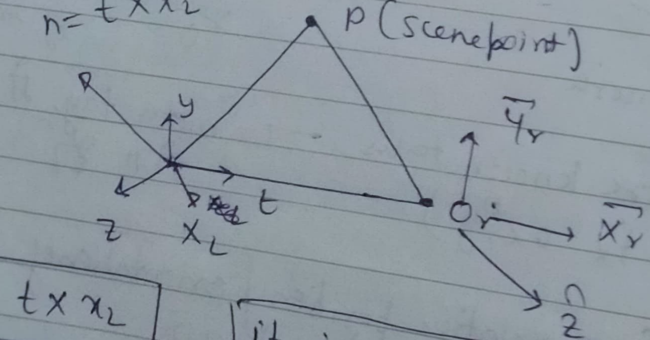


→ Camera matrix(k) is known to us
→ Find the small number of reliable correspoing points
  ↳ atleast (8) using ↳ SIFT
                              ↳ hand picked
→ Find Relative Camera
→ Position t and orentation R.

Find dense correspondence (map)
↳ compute Depth using Triangulation

$\boxed{\text{Epipolar Geometry}}$ →

↳ $\boxed{\text{Epipole}}$ → Image point of origin/pinhole of one
  ↳ camera as viewed by the other camera.



↳ $\boxed{n = t \times x_2}$ $\boxed{\text{it is normal to } x_2}$ → using RHT

therefore → ↳ $\boxed{x_L \cdot (t \times x_2) = 0}$ ↳ Epipolar constraint we don't have x_L and x_R but we have corresponding points in image coordinates

Rewriting in terms of image coordinates

$$[u_L \; v_L \; 1] \underline{z_L \, K^{-1^T}} \begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} K_r^{-1} z_r \begin{bmatrix} u_r \\ v_r \\ 1 \end{bmatrix} = 0$$

$\boxed{z_l \text{ can't be } 0 \quad z_r}$ → we can ignore this

This 3x3 matrix is also called fundamental matrix

$\boxed{E = k_l^T F K_r}$ $\boxed{E = T_x R}$

2025/03/02 21:40

$$[x_L \; y_L \; z_L] \underbrace{\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}}_{\text{essential matrix}} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

$\boxed{E = T_x R}$ — याद करें।

Longuet Huggin — 1981 की theorem है।

you can decompose it in $\underline{T_x}$ and R

$$\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix} = \begin{bmatrix} 0 & -t_z & t_y \\ t_z & 0 & -t_x \\ -t_y & t_x & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

skew

skew

and R → orthonormal

$\boxed{(AA^T) = 0}$

it is possible to "decouple T_x and R' from their product using Single Value Decomposition"

$\boxed{x_L^T E x_r = 0}$ — Epipolar c.

$$[x_L \; y_L \; z_L] \underbrace{\begin{bmatrix} e_{11} & e_{12} & e_{13} \\ e_{21} & e_{22} & e_{23} \\ e_{31} & e_{32} & e_{33} \end{bmatrix}}_{\text{essential matrix}} \begin{bmatrix} x_r \\ y_r \\ z_r \end{bmatrix} = 0$$

3D position in left coordinates

3D position in right camera