3D Vision

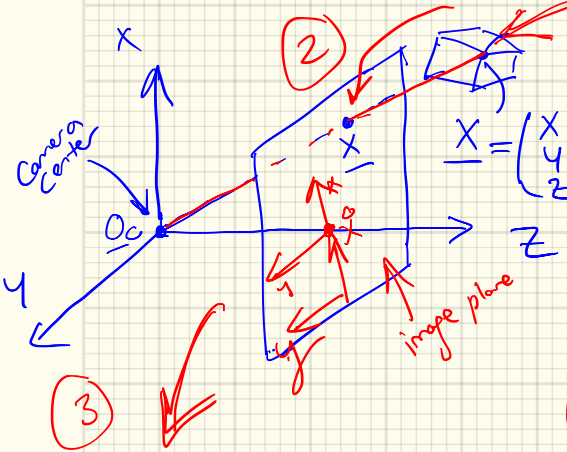
BLG634E

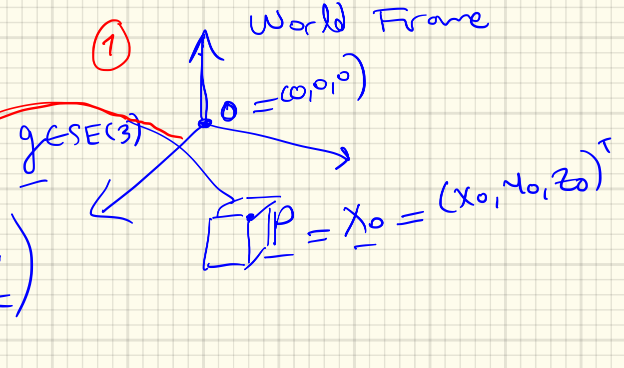
06.04.2021

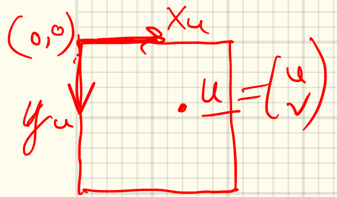
Gözde ÜNAL

28.03.2022

Camera Parameters







1. World to Camera Frame Coord. Xform
2. Projection of 3D Coordinates into 2D image Coord
3. Coordinate transform btw 2D metric (normalized) image coordinates $\underline{x}$ and 2D pixel coord. $\underline{u}$.

⓪ Move to homogeneous coord.

$$\begin{bmatrix} x_0 \\ y_0 \\ z_0 \end{bmatrix} \rightarrow \begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}$$

①

$$\underline{X}_0 \rightarrow \underline{X} = \bar{g}\, \underline{X}_0$$

$$\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} \underline{\underline{R}}_{3\times 3} & \underline{T}_{3\times 1} \\ \underline{0}^T_{1\times 3} & 1 \end{bmatrix}_{4\times 4}}_{\bar{g} \in SE(3) \rightarrow (\underline{\underline{R}}, \underline{T})} \begin{bmatrix} X_0 \\ Y_0 \\ Z_0 \\ 1 \end{bmatrix}$$

$$: \quad \underline{\underline{R}} \in SO(3), \quad \underline{T} \in \mathbb{R}^3$$

3D World to Camera Transformation

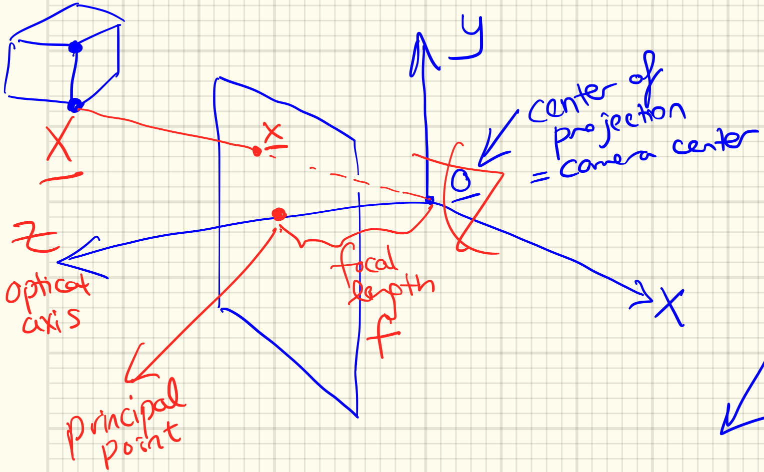
② Adapting pinhole camera model: Project $\underline{X}$ onto the image plane at the point $\underline{x} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix}$.

Recall perspective projection

$$\underline{x} = \begin{bmatrix} x \\ y \end{bmatrix} = \frac{f}{Z}\begin{bmatrix} X \\ Y \end{bmatrix}$$

↳ convert to homogeneous coord.

$$\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \frac{1}{Z}\begin{bmatrix} fX \\ fY \\ Z \end{bmatrix}$$



Z: depth, unknown

write Z as as a scalar

$$Z\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = \begin{bmatrix} f & 0 & 0 & 0 \\ 0 & f & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}_{3\times4}\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix}$$

$$\Rightarrow \quad \underbrace{\lambda \, \underline{x}}_{\begin{bmatrix} x \\ y \\ 1 \end{bmatrix}} = \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\underline{\underline{K}}_f \in \mathbb{R}^{3\times 3}} \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}}_{\underline{\underline{\Pi}}_0 \in \mathbb{R}^{3\times 4}} \underline{X} \rightarrow \begin{bmatrix} x \\ y \\ z \\ 1 \end{bmatrix}$$

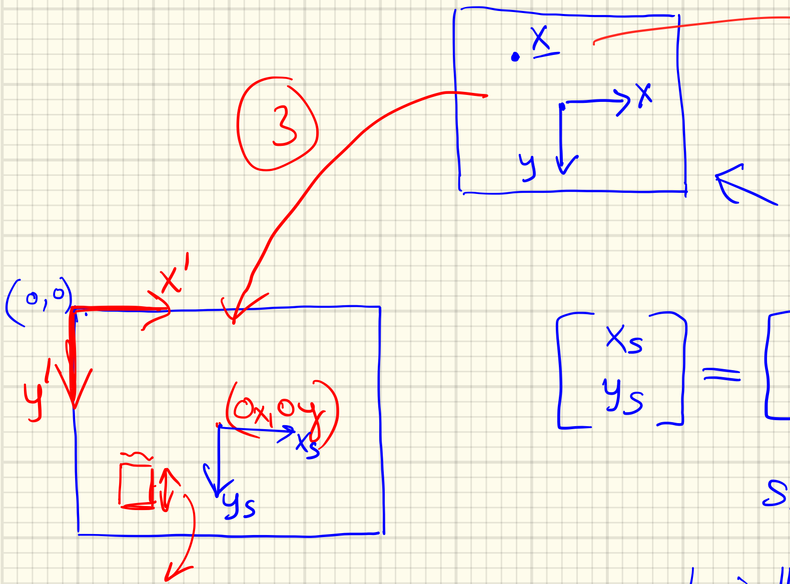
Π_0 : standard (canonical) projection matrix

$\mathbb{P}^3 \rightarrow \mathbb{P}^2$

To summarize :

$$\lambda \, \underline{x} = \underline{\underline{K}}_f \, \underline{\underline{\Pi}}_0 \, \underline{X} = \underline{\underline{K}}_f \, \underline{\underline{\Pi}}_0 \, \bar{g} \, \underline{X}_0$$

3. W/ a digital camera, measurements are typically provided in pixels w/ origin of image coordinate frame typically in the upper left corner of the image.



Metric Coordinates

(Normalized image coordinates)

$$\begin{bmatrix} x_s \\ y_s \end{bmatrix} = \begin{bmatrix} S_x & 0 \\ 0 & S_y \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$S_x, S_y \propto \frac{1}{mm}$, relate to size of pixels

typically $S_x = S_y$

$\begin{bmatrix} o_x \\ o_y \end{bmatrix}$: principal point

Translate the origin to upper left corner of the image.

$$x' = x_s + o_x$$
$$y' = y_s + o_y$$

(o_x, o_y) : coord. in pixels of the principal point on the image plane.

Now all pixel coordinates are POSITIVE.

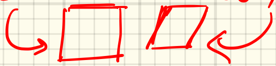
Actual (digital) image coord

$$\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix} = \underbrace{\begin{bmatrix} s_x & \boxed{s_\theta} & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{K_s \in \mathbb{R}^{3\times 3}} \begin{bmatrix} x \\ y \\ 1 \end{bmatrix}$$

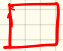
another linear transform

$\boxed{s_\theta : \text{skew factor}}$

$s_\theta = 0$ $\quad$ $s_\theta \neq 0$



When $s_x = s_y$



square

$s_\theta \neq 0$ when pixels are not rectangular

Usually $s_\theta = 0$
(assumed)

Now combine all the models

$$\lambda \underbrace{\begin{bmatrix} x' \\ y' \\ 1 \end{bmatrix}}_{\text{pixel coordinates}} = \underbrace{\underbrace{\begin{bmatrix} s_x & s_\theta & o_x \\ 0 & s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}}_{\underline{\underline{K}}_s} \underbrace{\begin{bmatrix} f & 0 & 0 \\ 0 & f & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\underline{\underline{K}}_f}} \underbrace{\begin{bmatrix} \underline{\underline{I}}_{3\times 3} & \begin{matrix} 0 \\ 0 \\ 0 \end{matrix} \end{bmatrix}}_{\underline{\underline{\Pi}}^0_{3\times 4}} \underbrace{\begin{bmatrix} \underline{\underline{R}} & \underline{T} \\ \underline{0}^T & 1 \end{bmatrix}}_{\overline{g}_{4\times 4}} \underbrace{\begin{bmatrix} x_0 \\ y_0 \\ z_0 \\ 1 \end{bmatrix}}_{\text{world point}}$$

$$\underline{\underline{K}} \triangleq \underline{\underline{K}}_s \underline{\underline{K}}_f$$

Π_0 : perspective Projection matric

$\underline{\underline{K}}$: intrinsic (internal) camera matrix :

depends on $\boxed{\text{intrinsic camera parameters}}$:

f : focal length

s_x, s_y, s_θ : scaling factors $\left(\frac{1}{mm}\right)$ & a skew $\longrightarrow \sim$

o_x, o_y : center offsets

$\overline{g}$: extrinsic camera parameters (R, T) (external)

$$\underline{\underline{K}} = \begin{bmatrix} f s_x & f s_\theta & o_x \\ 0 & f s_y & o_y \\ 0 & 0 & 1 \end{bmatrix}, \quad \lambda \underline{x}' = \underline{\underline{K}} \, \underline{\underline{\Pi}}_0 \, \underline{X}$$

$\underline{\underline{K}}$: upper triangular matrix : When the calibration matrix $\underline{\underline{K}}$ is known, the calibrated coord. $\underline{x}$ can be obtained from $\underline{x}'$ by:

$$\lambda \underline{x} = \lambda \underline{\underline{K}}^{-1} \underline{x}'$$

The problem of estimating $\underline{\underline{K}}$'s parameters is called intrinsic camera calibration.

Extrinsic Camera Calibration: Estimate $\bar{g}$ parameters

Defines geometric relation between the world point $\underline{X}_0$ & camera point $\underline{X}$.

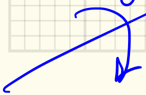
The overall model: $\lambda \underline{x}' = \underline{\underline{K}}\,\underline{\underline{\Pi}}_0\,\underline{X} = \underline{\underline{K}}\,\underline{\Pi}_0\,\bar{g}\,\underline{X}_0$
for pinhole camera projection

$$\lambda \underline{x'} = \underline{\underline{K}}\,\underline{\Pi}_0\,\bar{g}\,\underline{X}_0$$

Camera Calibration: 11 parameters to be estimated.

Intrinsics: $f\,,\ s_x,\ s_y\,,\ o_x,\ o_y$ $\quad (s_\theta) \quad \rightarrow 5$

Extrinsics: $\underline{\underline{R}}\,,\ \underline{\underline{T}} \quad \rightarrow 6$

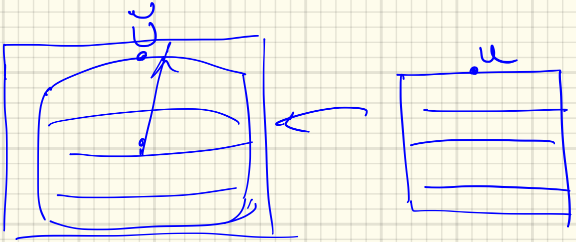


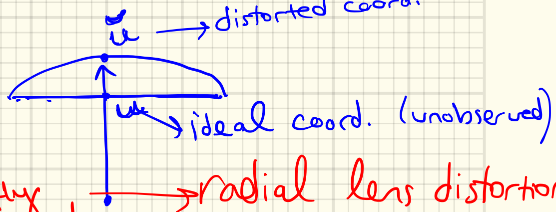
$\theta \rightarrow \underline{r} \leftarrow \begin{bmatrix} r_1 \\ r_2 \\ r_3 \end{bmatrix} \rightarrow 3 \qquad \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix}$

④ Lens Distortion Camera Parameters :

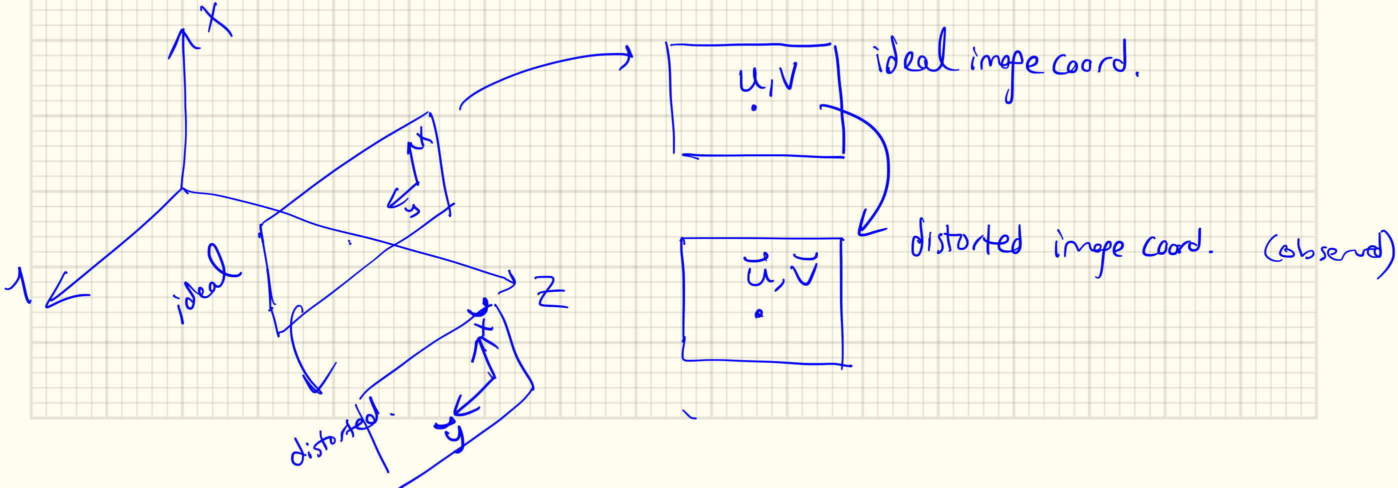
Optics of the lens introduces image distortions that show straight lines as curved lines.

effect





mostly adopted → radial lens distortion model :



$$\breve{x} = x\left(1 + k_1 r^2 + k_2 r^4\right) \qquad r^2 = x^2 + y^2$$

$$\breve{y} = y\left(1 + k_1 r^2 + k_2 r^4\right)$$

$\Rightarrow k_1, k_2$: (Radial) Lens Distortion Parameters

Write it in (u,v) coord.

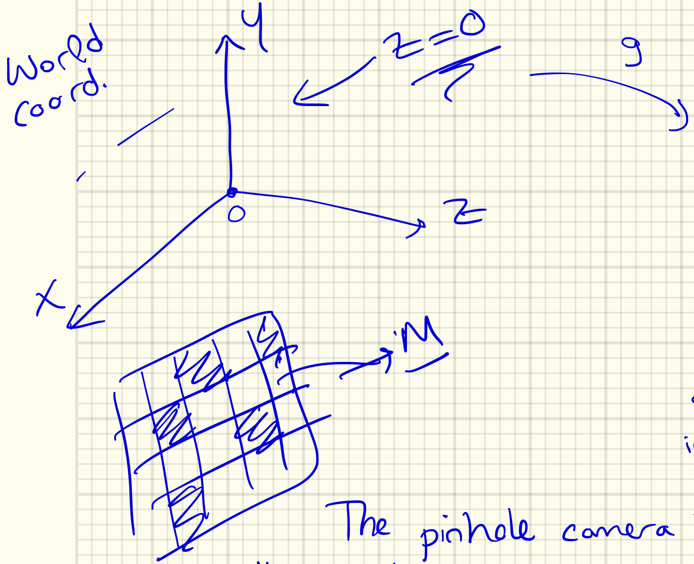
Radial Lens Distortion Model:

$$\breve{u} = u + (u - u_0)\left[k_1(x^2+y^2) + k_2(x^2+y^2)^2\right]$$

$$\breve{v} = v + (v - v_0)\left[k_1(x^2+y^2) + k_2(x^2+y^2)\right]$$

Lens Distortion Calibration $\Rightarrow$ Estimate k_1 & k_2.

Zhang's Camera Calibration: (Z. Zhang Technical Report)



Notation:

$$\underline{M} = \begin{pmatrix} X \\ Y \\ Z \end{pmatrix} \Rightarrow \tilde{M} = \begin{pmatrix} X \\ Y \\ Z \\ 1 \end{pmatrix}$$

3D model point

$$\underline{m} = \begin{pmatrix} u \\ v \end{pmatrix} \longrightarrow \tilde{m} = \begin{pmatrix} u \\ v \\ 1 \end{pmatrix}$$

2D image point

The pinhole camera is modeled by linking 3D point M & its projection m :

$$s\,\tilde{m} = \underline{\underline{A}}\;\underbrace{\begin{bmatrix} \underline{\underline{R}} & \underline{t} \end{bmatrix}}_{\text{extrinsic parameters}}$$

intrinsic camera matrix $\underline{\underline{A}} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix}$

α, β: scale factors

γ: skewness of the image axes.

$\tilde{u}$.

(u_0, v_0) : coord. of the principal point.

$\rightarrow$ Homography btw the $\overset{\text{calibration object}}{\text{model}}$ plane & its image $\qquad \boxed{\lambda\, \underline{x}' = \underline{K}\, \boxed{\Pi_0\, \bar{g}}\, X_0}$

Assume model plane is on $\underline{Z=0}$ of the world coord. frame.

$$\underbrace{s\begin{bmatrix} u \\ v \\ 1 \end{bmatrix}}_{\text{2D}} = \underline{\underline{A}}\, \overbrace{\begin{bmatrix} \underline{r_1} & \underline{r_2} & \underline{r_3} & \underline{t} \end{bmatrix}_{3\times 4}}^{= \Pi_0 \bar{g}} \underbrace{\begin{bmatrix} X \\ Y \\ 0 \\ 1 \end{bmatrix}}_{\text{3D}} \quad \leftarrow \text{assumption}$$

$$s\,\tilde{m} = \underbrace{\underline{\underline{A}} \begin{bmatrix} \underline{r_1} & \underline{r_2} & t \end{bmatrix}}_{\underline{\underline{H}}\ 3\times 3} \begin{bmatrix} X \\ Y \\ 1 \end{bmatrix} \longrightarrow$$

$\underline{\underline{H}}$ 3x3 homography defined upto a scalar

$$s\,\tilde{m} = \underline{\underline{H}}\,\tilde{M}$$

$\rightarrow$ Estimate $\underline{\underline{H}}$ first. Given n model points $\{M_i\}_{i=1}^{n} \longleftrightarrow \{m_i\}_{i=1}^{n}$

(Appendix A)

model points $\longleftrightarrow$ image points.

Sec 2.3

$$\begin{bmatrix} \underline{h_1} & \underline{h_2} & \underline{h_3} \end{bmatrix} = \underset{\sim}{\lambda}\, \underline{\underline{A}} \begin{bmatrix} \underline{r_1} & \underline{r_2} & \underline{t} \end{bmatrix}$$

$$\bigstar \quad \boxed{\begin{pmatrix} u \\ v \end{pmatrix} = \hat{m}_i = \frac{1}{\underline{h}_3^T \underline{M}_i} \begin{pmatrix} \underline{h}_1^T \underline{M}_i \\ \underline{h}_2^T \underline{M}_i \end{pmatrix}} \qquad \Leftarrow \quad \hat{m} = \underline{\underline{H}}\, \hat{M}$$

$$\underline{\underline{H}} = [\underline{h}_1 \quad \underline{h}_2 \quad \underline{h}_3]$$

h_i : ith row of $\underline{\underline{H}}$

$$\min_{\underline{\underline{H}}} \sum_i \| m_i - \hat{m}_i \|^2$$

m_i: measured image points

$\hat{m}_i$: projection of 3D model points M_i onto the image.

Nonlinear optimization problem: Uses Levenberg-Marquardt (LM) Algorithm (in minpack) (matlab "lsqnonlin").

→ Get an initial condition (guess) for $\underline{h}$:

Note $\hat{m} = \begin{pmatrix} u \\ v \end{pmatrix}$ Rewrite ($\bigstar$) →

$$\left.\begin{aligned} u\, \underline{h}_3^T \tilde{M}_i &= h_1^T \tilde{M}_i \\ v\, h_3^T \tilde{M}_i &= h_2^T \tilde{M}_i \end{aligned}\right\}$$

$$\underbrace{\begin{bmatrix} \tilde{M}^T & \underline{0}^T & -u\tilde{M}^T \\ \underline{0}^T & \tilde{M}^T & -v\tilde{M}^T \end{bmatrix}_{2n\times 9}}_{\underline{\underline{L}}_{2n\times 9}} \underbrace{\begin{bmatrix} \underline{h}_1 \\ \underline{h}_2 \\ \underline{h}_3 \end{bmatrix}_{9\times 1}}_{\underline{h}_{9\times 1}} = \underline{0}$$

$$\boxed{\underline{\underline{L}}\, \underline{h} = 0} \Longrightarrow$$ Solution to this homogeneous system of equations is the eigen vector of $\underline{\underline{L}}^T \underline{\underline{L}}$ corresp. to the smallest eigenvalue.

$\rightarrow \underline{h}$ ✓

Note

$$\underline{\underline{L}}\,\underline{h} = \begin{bmatrix} x & y & 1 & 0 & 0 & 0 & -u[x\ y\ 1] \\ 0 & 0 & 0 & x & y & 1 & -v[x\ y\ 1] \end{bmatrix} \begin{bmatrix} h_{11} \\ h_{12} \\ h_{13} \\ h_{21} \\ h_{22} \\ h_{23} \\ h_{31} \\ h_{32} \\ h_{33} \end{bmatrix}$$

$$\underline{\underline{H}} = [\underline{h_1}\ \ \underline{h_2}\ \ \underline{h_3}] \checkmark$$

Use constraints $[\underline{h_1}\ \ \underline{h_2}\ \ \underline{h_3}] = \lambda\,\underline{\underline{A}}\,[r_1\ \ r_2\ \ r_3]$

$$\underline{r_1} = \underline{\underline{A}}^{-1}\underline{h_1}$$

$$\underline{r_2} = \underline{\underline{A}}^{-1}\underline{h_2}$$

(1) $\underline{r_1}^T \underline{r_2} = 0$

(2) $r_1^T r_1 = r_2^T r_2 = I$

(1) $\underline{h_1}^T \underline{\underline{A}}^{-T}\underline{\underline{A}}^{-1}\underline{h_2} = 0$

(2) $h_1^T \underbrace{A^{-T} A^{-1}}_{B} h_1 = h_2^T \underbrace{A^{-T} A^{-1}}_{\underline{\underline{B}}} h_2$

$$\underline{\underline{B}} = \begin{bmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{23} & B_{23} \\ B_{31} & B_{32} & B_{33} \end{bmatrix}$$

$$\underline{\underline{A}} = \begin{bmatrix} \alpha & \gamma & u_0 \\ 0 & \beta & v_0 \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \underline{\underline{A}}^{-T} \underline{\underline{A}}^{-1} = \underline{\underline{B}}$$

$$\underline{\underline{B}} = \begin{bmatrix} \frac{1}{\alpha^2} & -\frac{\gamma}{\alpha^2 \beta} & \\ \vdots & & \\ & & \end{bmatrix}_{3\times 3}$$

symmetric

Eq. 5 $\Rightarrow$

$$\underline{b} = \begin{bmatrix} B_{11} \\ B_{12} \\ \vdots \\ B_{33} \end{bmatrix}_{6\times 1}.$$

① × ②

$$\Rightarrow \underbrace{\underline{h_i}^T \underline{\underline{B}} \, h_j = V_{ij}^T \underline{b}}$$

$$V_{ij} = \begin{bmatrix} h_{i1} h_{j1} \\ h_{i1} h_{j2} + h_{i2} h_{j1} \\ \vdots \end{bmatrix}_{6\times 1}$$

$$\left.\begin{aligned} &① \quad h_1^T B h_2 = 0 \\ &② \quad h_1^T B h_1 - h_2^T B h_2 = 0 \end{aligned}\right\} \Rightarrow \underset{2\times 6}{\underline{V}} \; \underset{6\times 1}{\underline{b}} = \underline{0} \Rightarrow \begin{bmatrix} V_{12}^T \\ (V_{11} - V_{22})^T \end{bmatrix} \underline{b} = 0$$

h_i

$\underline{\underline{B}}$

$\Rightarrow$ Once $\underline{b}$ is estimated; α, β, γ, u_0, v_0 are estimated ✓

$\Rightarrow$ Now, all camera intrinsic matrix $\underline{\underline{A}}$ are known ✓

$\Rightarrow$ Now, extrinsics can be calculated.

$\Rightarrow$ ★ $\underline{r_1} = \lambda \underline{\underline{A}}^{-1} \underline{h_1}$, $\underline{r_2} = \lambda \underline{\underline{A}}^{-1} \underline{h_2}$, $\underline{r_3} = \underline{r_1} \times \underline{r_2}$

$$\Rightarrow \lambda = \frac{1}{\|A^{-1} h_1\|} \qquad \lambda$$

$$\underline{t} = \lambda \underline{\underline{A}}^{-1} \underline{h_3}$$

At this point, we have initial estimates for all intrinsic & extrinsic parameters ✓

Given n images of the model plane & m points on the model plane

Set up the optimization:

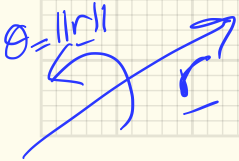
$$\min \sum_{\underset{n \text{ images}}{i=1}}^{n} \sum_{\underset{m \text{ points}}{j=1}}^{m} \| \underbrace{m_{ij}}_{\text{measurement } j \text{ on image plane } i} - \underbrace{\hat{m}(\underline{\underline{A}}, \underline{\underline{R}}_i, \underline{t}_i, \underline{M}_j)}_{\text{projection of point } M_j \text{ onto image } i} \|^2$$

Nonlinear optimization problem (eg. solve by LM Algo).

Bundle Adjustment $\leftarrow$

Note: $\underline{\underline{R}}$ is parameterized by $\underline{r}$: a 3D axis vector

its magnitude is the rotation angle



Using Rodrigues formula $\underline{\underline{R}} \Longleftrightarrow \underline{r}$