24 November 2020 23:00

Problem	Structure (Scene geometry)	Motion (Camera parameters)	Measurements
F-Matrix estimation	Unknown	Estimate	2D-2D correspondences
Camera Calibration	known	Estimate	2D-3D correspondences.
Triangulation	Estimated	known	2D-3D correspondences
Stereo Rectification	estimated	known	2D-2D correspondences
PnP	known	estimated	2D-3D correspondences
Bundle Adjustment	estimated	estimated	2D-3D correspondences

(a) We have,

$$l = F^T x'$$

where x' is the point of second image.

l is the epipolar line of first image plane.

Since we have

$$x'^T F x = 0$$

$\forall$ x in first & x' in second image planes,

we have,

$$x'^T F x = 0$$

$$\therefore \; x^T F^T x' = 0$$

Since $x = e$, lies on the epipolar line given by $F^T x'$,

we have

$$e^T l = 0$$

or

$$e^T (F^T x') = 0$$

$$e^T (F^T x') = 0$$

or

$$(e^T F^T)\, x' = 0$$

or

$$(Fe)^T x' = 0$$

$$\therefore \boxed{Fe = 0}$$

'e' is right null vector.

(b) In case of pure rotation,

$$P = K[I \mid 0] \quad P' = K[R \mid 0]$$

$$P^+ = \begin{bmatrix} K^{-1} \\ 0^T \end{bmatrix} \quad C = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Recall,

$$F = [e']_\times P' P^+$$

[epipolar line joining e' to $P'P^+$].

e' to $P'P^+]$.

$\therefore F = [\cancel{P}C]_\times P'P^+$

$= [K\ 0]_\times KRK^{-1}$

$= KRK^{-1}$

(c) If F takes images i1 and i2, and for any point in i2, we have a line in i1 such that

$$\underline{l_2 = F x_1}.$$

F can be decomposed as

$$F = [e_2]_\times H_\pi$$

where e_2 is second image's epipole and

and H_π maps from image 1 to image 2.

Since, for any point x_2 lies on l_2, we have

$$x_2^T l_2 = 0$$

or

$$x_2^T F x_1 = 0$$

$$\therefore \quad \left((F x_1)^T x_2 \right)^T = 0$$

$$\therefore \quad \boxed{x_1^T F^T x_2 = 0}$$

$$\therefore \quad \boxed{x_1^T \; F^T x_2 = 0}$$

Between image 2 & image 1, we have F^T as our fundamental matrix, as for any x_2, it has any epipolar line $(F^T x_2)$ which has a set of points (x_1) corresponding lying on it. Indeed, we can also think of it as from image 2 to image 1, $(H_\pi)^T$ will do the job,

and thus

$$F' = [e_2]_\times^T (H_\pi)^T = F^T.$$

(d) Essential matrix takes into account the normalized coordinates, that is, for every x, we have

$$\hat{x} = K^{-1} x$$

so that

$$\boxed{\hat{x} = [R \mid t] X}$$

Such a case, would have essential matrix derivation as:—

$$P = [I \mid 0] \quad P' = [R \mid t]$$

$$P^{+} = \begin{bmatrix} I \\ 0^T \end{bmatrix} \quad C = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$E = K' = [P'C]_{\times} [P'P^{+}]$$

$$= [0 \ \ t]_{\times} [R \ 0]$$

$$= \underline{[t]_{\times} R}$$

equation for essential matrix is

$$- \hat{x}'^{T} E \hat{x} = 0$$

instead of

$$x'^{T} F x = 0$$

in case of fundamental matrix.

now, substitute $\hat{x}'$ and $\hat{x}$ above -

$$x'^{T} \underline{K'^{-T} E K^{-1}} x = 0$$

instead of

$$x'^{T} \underline{F} x = 0$$

$$\therefore \quad K'^{-T} E K^{-1} = F$$

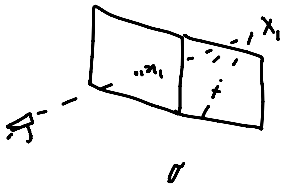
$\therefore \quad \boxed{E = K'^{T} F K}$

Essential matrix has only 5 degrees of freedom compared to 7 degrees of freedom.

DLT (Direct Linear Transform)	Zhang's Method
- 11 parameters using atleast 6 points for complete calibration.	- 8 parameters, need 4 points. [each point has 2 observations]
- Estimate 3 x 4 matrix	- Estimate 3x3 matrix.
- Do not ignore (Z_i) $a_{x_i}^T = (-X_i, -Y_i, -Z_i, -1, 0, 0, 0, 0, x_iX_i, x_iY_i, x_iZ_i, x_i)$ $a_{y_i}^T = (0, 0, 0, 0, -X_i, -Y_i, -Z_i, -1, -y_iX_i, -y_iY_i, -y_iZ_i, y_i)$	- Ignore (Z_i) → $a_{x_i}^T = (-X_i, -Y_i, -1, 0, 0, 0, x_iX_i, x_iY_i, x_i)$ $a_{y_i}^T = (0, 0, 0, -X_i, -Y_i, -1, y_iX_i, y_iY_i, y_i)$
- can handle non-planar data. [those having different Z_i]	- requires planar target. [like printed checkerboard]
- solved by solving $Vx=b$ linear equation with 11 dof 1 constraint	- solved using SVD of M. [take by stacking 2 ...

linear equation with ll6lF + constraint.	M. [taking stacking 2I vectors].

(a)



We have

$$x_1 = K[I \mid 0]X = KX$$

$$x_2 = K[R \mid 0]X$$

$\therefore$ we have

$$x_2 = K[R \mid 0]X = K[R \mid 0]K^{-1}x_1$$

$$= \underline{\underline{KRK^{-1}}}x_1$$

Hence x_2 is related to x_1 by the homography matrix

$$\boxed{x_2 = Hx_1}$$

where $\underline{\underline{H = KRK^{-1}}}$.

(b) We have, if R denotes rotation by θ, then R^2 denotes rotation by $\underline{\underline{2\theta}}$.

Proof:
Indeed, we have rotation by 2θ, involve rotating by 'θ' twice.

$$\therefore \quad R(2\theta) = R(\theta)R(\theta)$$

$$= \underline{\underline{R^2}}$$

Now, for 2θ we would have

$$x_1 = K[I \mid 0]X$$

$x_1 = K[I \mid 0]X$

$x_2 = K[R^2 \mid 0]X.$

$\therefore \underline{x_2 = KR^2K^{-1}x_1}$

$\therefore \underline{x_2 = KR \cdot R \cdot K^{-1} x_1}.$

$\therefore \underline{x_2 = KR(K^{-1}K)RK^{-1}x_1}$

$\therefore x_2 = (KRK^{-1}).(KRK^{-1})x_1$ -- {matrix multiplication is associative}.

$= H \cdot H x_1$

$= H^2 x_1$

Thus

$H^2 = (KRK^{-1})(KRK^{-1})$

$= KR^2K^{-1}$ as proved above.

(a) Given P [camera matrix],

we have

$$P = KR[I_3 | -X_0]$$

[here K, R are projected that is $\hat{K}, \hat{R}$].

$$= [\underset{3\times 3}{\hat{H}_\infty} \,|\, \underset{3\times 1}{\hat{h}}].$$

where

$$\hat{H}_\infty = KR$$

and

$$\hat{h} = -KRX_0$$

$\therefore$ projection center,

$$\boxed{X_0 = -\hat{H}_\infty^{-1} \hat{h}}$$

We use QR decomposition over here,

any square matrix A can be decomposed into Q and R:-

$$\underline{A = QR}$$

Q is an orthogonal matrix

R is an upper triangular matrix.

Now, we have

$$\hat{H}_\infty = KR$$

$$\therefore (\hat{H}_\infty)^{-1} = R^{-1} K^{-1}$$

$$\therefore (\hat{H}_\infty)^{-1} = R^{-1} K^{-1}$$
$$= \underline{\underline{R^T K^{-1}}}.$$

$\therefore$ QR decomposition of $(\hat{H}_\infty)^{-1}$ should give $\underline{R^T}$ and $\underline{K^{-1}}$.

$$(\hat{H}_\infty)^{-1} = QR'$$

we have

$$\boxed{\begin{aligned} R &= Q^T \\ K &= (R')^{-1} \end{aligned}}$$

we can normalize K,

$$\boxed{K = \frac{1}{\hat{K}_{33}} K}$$

for homogenity.

(b)

(c) If there isn't known about the calibration of 2 cameras, the reconstruction falls into what we call projective reconstruction.

for any image point project

$$x_i = P_j X_i$$

for a camera matrix P_j,

we can apply a projective transformation H to each point X_i, →

$$\boxed{x_i = (P_j H^{-1})(H X_i)}$$

Thus without changing the image, we were able to apply a projective transformation, which is why we say that it has projective ambiguity.

The fundamental matrix 'F' derived from

$$\underline{x_i'^T F x_i = 0}.$$

the reconstructed 'P' from F are related by the a projective transformation $\underline{H}$: This may occur when there are 7 points

(b) Transformation from camera to image frame represents perspective transformation, that results in a planar coordinate on its image plane.

Note that depth of image plane in camera coordinate system is focal length, 'f', but here

coordinate system is focal length, 'f', but there is no depth in image plane coordinates.
Instead, we have a scaling factor, λ,

$$\boxed{\lambda x = KX}$$

This 'λ' is unknown if there is an ambiguity in calibration, fundamental matrix, R, t.

(a) It is possible to remove the ambiguity from the reconstruction if camera calibration matrix is known along with precisely knowing R, t.

Another way could be to have some "ground-control" points, some points using ground truth, for each point X_i, we can get the projective transformation H,

$$\underset{\substack{\text{(ground)}\\\text{truth}\\\text{point}}}{X_{Ei}} = HX_i$$

4th point

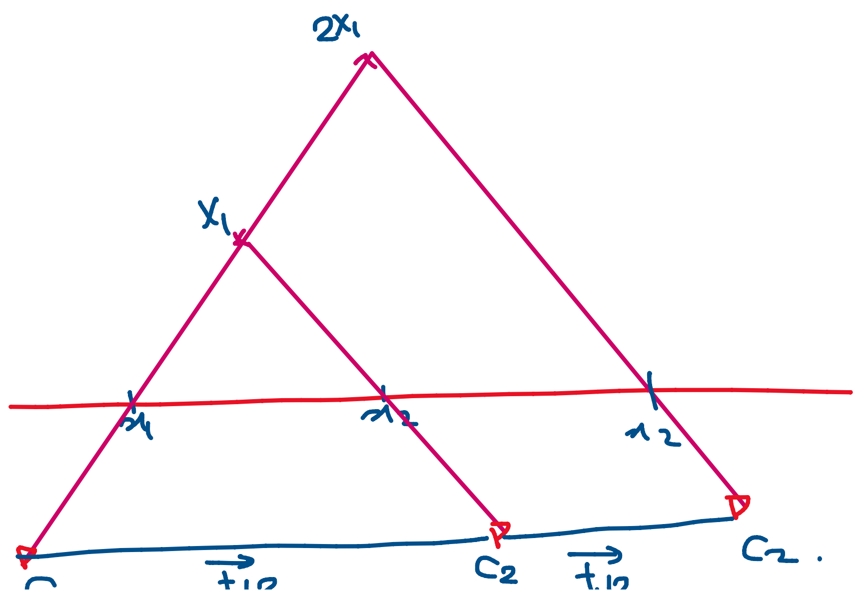
$$x_i = P H^{-1} X_{E i}$$

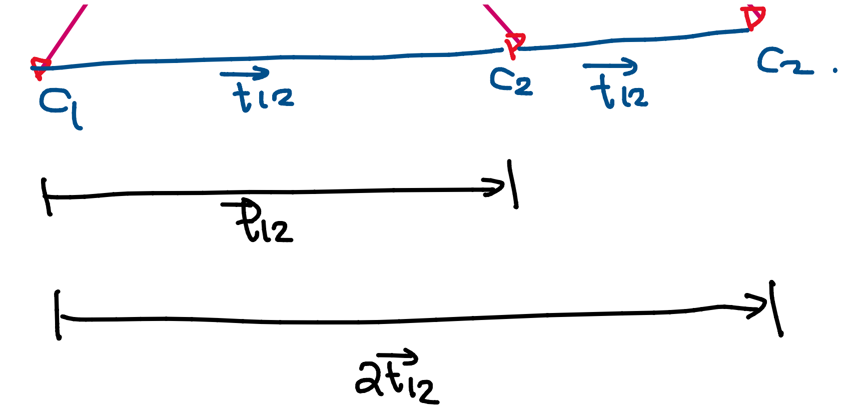
each point gives 3 linearly independent equation,
and H has 15 dof,
we must have $n \geq 5$ points
such that no 4 points are coplanar.

(c) Geometrically speaking,

1) the camera moving by 1 unit and object at X distance from camera.

2) the camera moving by 2 units and object at 2X distance from camera.





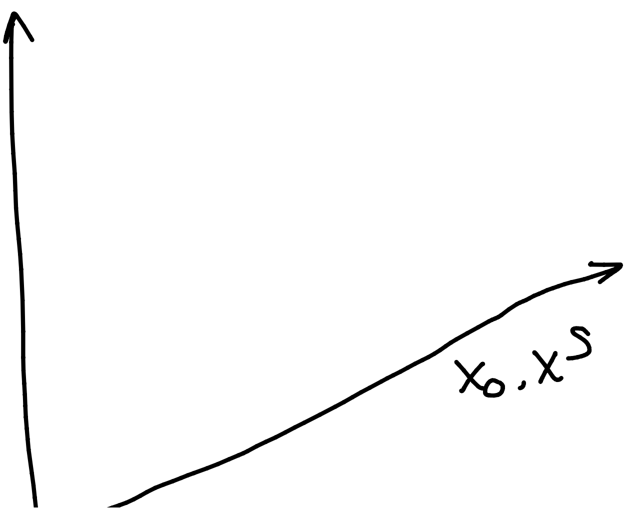
both have the same reconstruction.

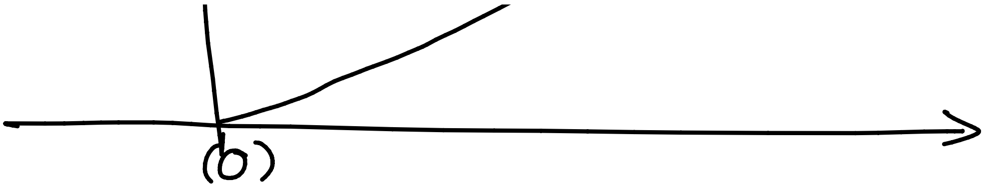
(d) <u>Mathematically</u>,

we will demonstrate that points have a scalability transform between them.

Given 2 frames S and O, we such that 1 unit in S frame is k units in O frame.

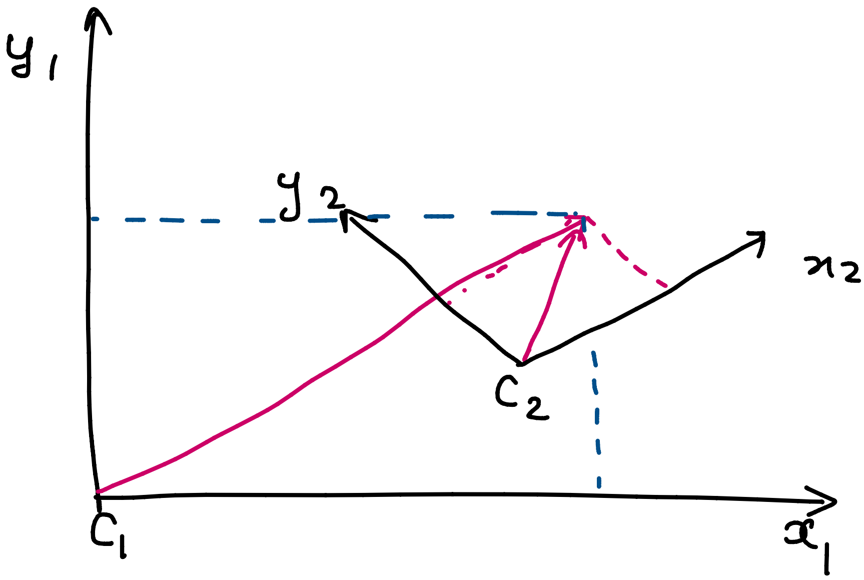
In that case, we have





$$X^0 = K X^s$$
$$= K [I \mid 0] X^s$$
$$=$$

the transformation between 2 coordinates

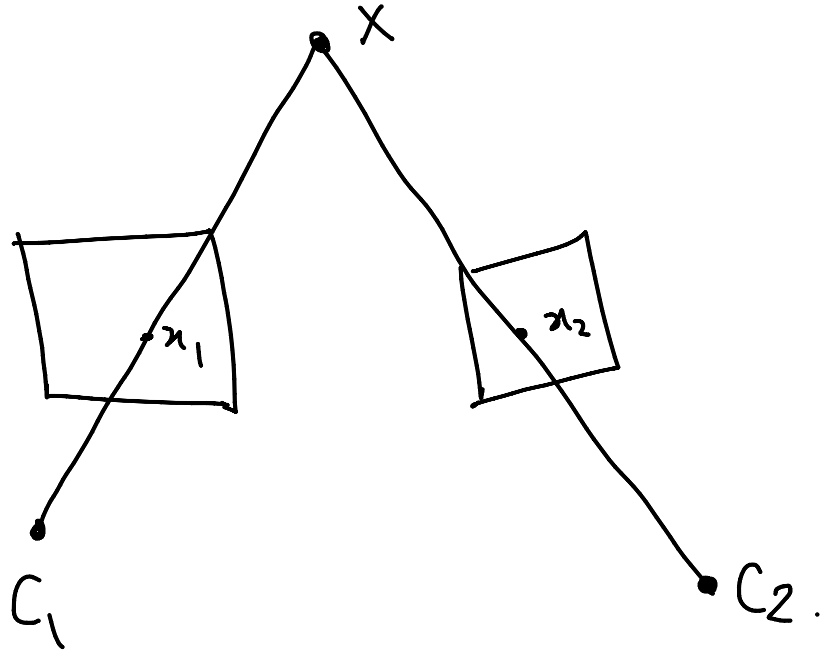


then we have,

$$X^1 = \begin{bmatrix} K R & t \\ 0 & 1 \end{bmatrix} X^2$$

let's look at the problem of 2 views again.

Let's look at the problem of 2 views again, say, if we had 2 camera C_1 and C_2, but we wanted to know the scaling factor 'λ' correctly.



Say we have 2 frames C_1' and C_2' such that C_1 coincides with C_1' and C_2 coincides with C_2'.

The point X in C_1' would be $\lambda X_1'$ due to scaling. Thus, we have

$$x_1 = K[I \mid 0]X$$
$$= K\left[\frac{I}{h} \,\middle|\, 0\right] hX$$

Now let camera C_2 be at R_{21} rotation t_{21} translation wrt C_1.

ie.

$$x_2 = K[R_{21} \mid t_{21}]X$$

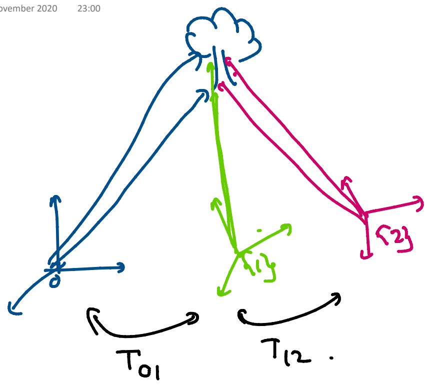
now, we know that C_2 is at t_{21} wrt C_1 and ht_{21} wrt $\underline{C_1'}$. Thus, the above equation cannot be distinguished from a camera C_2'.

$$\boxed{x_2 = K\left[\frac{1}{h} R_{21} \,\middle|\, t_{21}\right] hX}$$

6

24 November 2020 23:00

(a)



Cost function

$$\sum_{i=1}^{m} \sum_{j=1}^{n} \left\| \vec{X}_{ij} - \hat{T}_{i0}\vec{X}_{0p} \right\|_2^2$$

where

X_{ij} : j^{th} point in i^{th} frame,
X_{0j} : j^{th} point in 0^{th} frame

} direct measurements (say from LiDAR)

T_{i0} : homogenous transform matrix between i^{th} and 0^{th} observation.

m: no of frames/point clouds.

n: no of points in each frame/point cloud.

(b) We have

$$\underset{X_i, P_i}{\arg\min} \sum_{i=1}^{m} \sum_{j=1}^{n} \left\| \frac{P_{(1,2)i}\vec{X}_j}{P_{3i}\vec{X}_j} - \vec{x}_{ij} \right\|^2$$

P_i : i^{th} view's camera matrix

X_j : j^{th} point

x_{ij} : j^{th} pixel in i^{th} image.

(c) for ICP-Slam, we have Jacobian J as

$$
J = \left[\begin{array}{ccccc|c}
\multicolumn{5}{c|}{\overbrace{\hspace{6em}}^{\text{'m' times}}} & \overbrace{\hspace{6em}}^{\text{'n' times}} \\
J_{11} & 0 & - & - & - & \\
J_{12} & 0 & - & & - & \\
\vdots & \vdots & - & & - & \\
0 & 0 & - & & & \\
0 & J_{m2} & - & & & \\
\vdots & J_{2n2} & - & & - & \\
\vdots & 0 & - & & & \\
\vdots & \vdots & & & J_{mn} & \\
 & & & & J_{mn2} &
\end{array}\right]
\begin{bmatrix} \delta\xi_1 \\ \delta\xi_{2n} \\ \vdots \\ \delta\xi_n \\ \delta x_{01} \\ \vdots \\ \delta_{0n} \end{bmatrix}
$$

localization jacobian   mapping jacobian.

$(6m+3n)$ columns

&

$\underline{3mn \text{ rows}}$..

for monocular vision, we have

$(12m+3n)$ columns & $\underline{(2mn) \text{ rows}}$.

(d)

In case of ICP-Slam, we have

$12m+3n$ optimization variables,

$\vec{X}_{0j} \quad j = 1 \rightarrow n$

$\vec{T}_{0i} \quad i = 1 \rightarrow m$

In case of Monocular Slam, we are

minimizing over $P_1, P_2 \ldots P_M$

minimizing over $P_1, P_2 \dots P_M$

and $X_1, X_2 \dots X_N$.

which is only we's computing

$$\left[\frac{\partial r}{\partial P_1}, \frac{\partial r}{\partial P_2} \dots \frac{\partial r}{\partial P_M} \;\middle|\; \frac{\partial r}{\partial X_1}\; \frac{\partial r}{\partial X_2} \dots \frac{\partial r}{\partial X_N}. \right]$$

(e)

<u>Update step for ICP Slam:-</u>

We have our variables,

$$X_{oj}(n+1) = X_{oj}(n) - \delta X_{oj}$$

or, in general

$$\underline{X_j(n+1) = X_j(n) - \delta X_j}$$

[since T is a matrix, and not a vector, we use the update rule:-

$$T_{io}(n+1) = T_{io}(n) \cdot \exp([\delta \xi_{io}]_\times)$$

where ξ is an arbitrary axis vector

$$\mathcal{E}_i = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ v_1 \\ v_2 \\ v_3 \end{bmatrix}$$

$$J_{ij} = \left[\left(\frac{\delta r_{ij}}{\delta T_i} \frac{\partial T_i}{\partial \mathcal{E}_i} \right) \left(\frac{\delta r_{ij}}{\delta X_j} \right) \right]$$

using Gauss newton update here, we get :-

$$\begin{bmatrix} \delta \mathcal{E}_i \\ \delta x_j \end{bmatrix} = [J^T J]^{-1} \quad J^T \quad [\vec{X}_{ij} - \hat{R}_{io} \hat{x}_{oj} + \hat{t}_{io}]$$

from the update equation

$$\underline{\delta = -\alpha J^T r}$$

[Not the $\hat{R}_{io}$ & $\hat{t}_{io}$ can be obtained from $\hat{T}_{io}$].

Update step for monocular vision

we have

$\delta l c \longrightarrow$ update vector

$$J = \left[\begin{array}{cccc|cccc} J_{11} & 0 & 0 \cdots & 0 & J_{11s} & 0 & 0 & 0 \\ G_{11m} & & & & G_{11s} & 0 & 0 & - \\ & & & & 0 & J_{12s} & 0 & \\ J_{12m} & & & & & G_{12s} & & \\ G_{12n} & & & & 0 & & & J_{M s} \\ & & & & & & & G_{M s} \\ \hline 0 \cdot J_{21m} & & & & & & & \\ G_{21} - & & & & & & & \\ \hline \vdots & & & J_{M1m} & J_{M1s} & & & \end{array}\right] \quad \begin{bmatrix} \delta p_{111} \\ \vdots \\ \delta p_{1\,34} \\ \delta p_{2\,11} \\ \vdots \\ \delta p_{2,34} \\ \hline \delta x_1 \\ \delta y_1 \\ \delta z_1 \end{bmatrix}$$

$$
\begin{bmatrix}
 & & \vdots & J_{M1m} & J_{M1s} & \\
 & & \vdots & G_{M1m} & G_{M1s} & \\
 & & & & & J_{MNs} \\
 & & & & & G_{MNs}
\end{bmatrix}
\begin{bmatrix}
\delta x_1 \\ \delta y_1 \\ \delta z_1 \\ \vdots \\ \delta x_N \\ \delta y_N \\ \delta z_N
\end{bmatrix}
$$

where

$$J_{ijm} = \left[\frac{\partial f_{ij}}{\partial p_i}\right]$$

$$G_{ijm} = \left[\frac{\partial g_{ij}}{\partial p_i}\right]$$

$$J_{ijs} = \left[\frac{\partial f_{ij}}{\partial \vec{x}_j}\right]$$

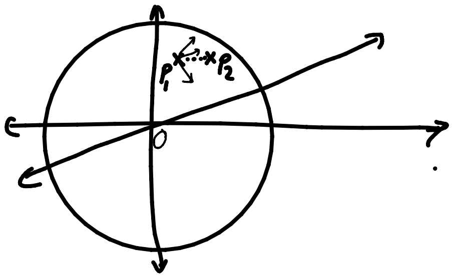
$$\vec{G}_{ijs} = \left[\frac{\partial g_{ij}}{\partial \vec{X}_j}\right].$$

$\Delta \vec{k}$ → update vector

$$\Delta k = -J^T r(k)$$

☐

24 November 2020 23:00



Assume $R = 6400km$
$= 6.4\times10^6 m$

We have,

$$T^{O}_{P_2} = T^{O}_{P_1} \cdot T^{P_1}_{P_2}.$$

we know,

$$T^{O}_{P_1} = \begin{bmatrix} R & t \\ \bar{0} & 1 \end{bmatrix}$$

P_1 has coordinates of

$$\left(\frac{\sqrt{3}R}{2}, R/2, R/2\right)$$

We know that latitude $= 30°$,
longitude $= 30°$,

$\therefore$ $R_z \Rightarrow 30°$, (ψ) $\quad R_x \Rightarrow 0°$ (ϕ)
$R_y \Rightarrow 30°$. (θ)

$\therefore$ Using ZYX, we have

$$R = R_z(\psi) R_y(\theta) R_x(\phi)$$

$$= \begin{bmatrix} \cos\theta\cos\psi & -\cos\phi\sin\psi + \sin\phi\sin\theta\cos\psi & \sin\phi\sin\psi + \cos\phi\sin\theta\cos\psi \\ \cos\theta\sin\psi & \cos\phi\cos\psi + \sin\phi\sin\theta\sin\psi & -\sin\phi\cos\psi + \cos\phi\sin\theta\sin\psi \\ -\sin\theta & \sin\phi\cos\theta & \cos\phi\cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} 3/4 & -1/2 & \sqrt{3}/4 \\ \sqrt{3}/4 & 7/8 & 1/4 \\ -1/2 & 0 & \sqrt{3}/2 \end{bmatrix}$$

Now, we will look at $T^{P_1}_{P_2}$,

we know $R=0$, $T = \begin{bmatrix} 0 \\ 500/\sqrt{2} \\ 500/\sqrt{2} \end{bmatrix}$

$$T^{0}_{P_2} = \begin{bmatrix} 3/4 & -1/2 & \sqrt{3}/4 & \frac{3R}{4} \\ \sqrt{3}/4 & 7/8 & 1/4 & \frac{\sqrt{3}R}{4} \\ -1/2 & 0 & \sqrt{3}/2 & R/2 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 500/\sqrt{2} \\ 0 & 0 & 0 & 500/\sqrt{2} \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

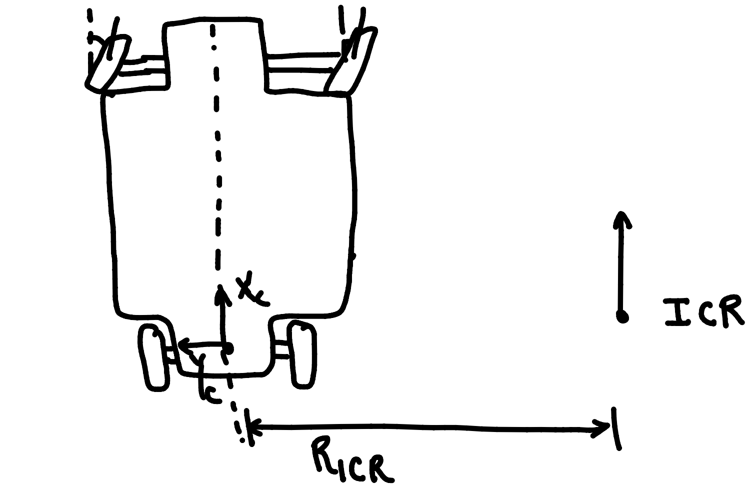
$$= \begin{bmatrix} 0 & 0 & 0 & 479976.31 \\ 0 & 0 & 0 & 2771594.56 \\ 0 & 0 & 0 & 3200306.18 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

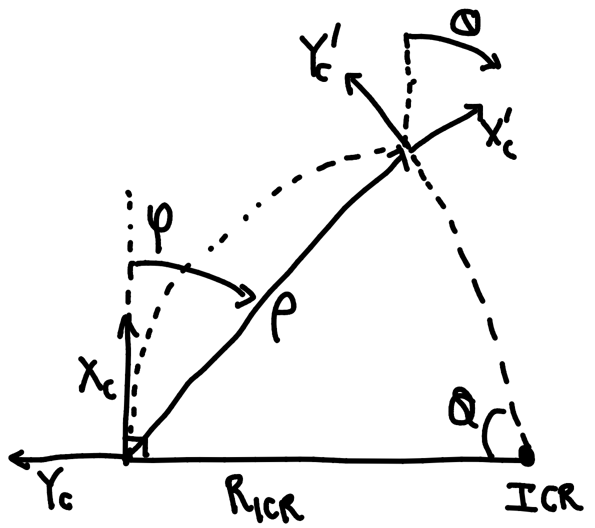
$\therefore$ we get the $P_2 = (479976.31, 2771594.56, 3200306.18)$

$\therefore$ we get the $r_2 = (479946.31, 2771194.16, 3200306.18)$

8

24 November 2020 23:03





a

We have

$$\rho = 2R_{ICR}\sin\theta/2 \quad \text{..(1)}$$

$$\varphi = \theta/2 \quad \text{----(2)}$$

Since car is non-holonomic, we assume the circular motion and come to the above conclusion.

For epipolar geometry, we have

$$E = [T]_x R$$

since R rotates by 'θ', we have

$$R = \begin{bmatrix} \cos\theta & +\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$T = \begin{bmatrix} p\cos\theta/2 \\ -p\sin\theta/2 \\ 0 \end{bmatrix}$$

$$E = [T]_x R$$

$$= \begin{bmatrix} 0 & 0 & -p\sin\theta/2 \\ 0 & 0 & -p\cos\theta/2 \\ +p\sin\theta/2 & p\cos\theta/2 & 0 \end{bmatrix} \begin{bmatrix} \cos\theta & +\sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & -p\sin\theta/2 \\ 0 & 0 & -p\cos\theta/2 \\ -p\sin\frac{\theta}{2} & +p\cos\theta/2 & 0 \end{bmatrix}$$

(b) Initially it looks like we have

(b) Initially it looks like we have

'θ' and R_{ICR}. as our parameters.

We take a pair of image points (p, p')

taken from the 2 orientations of the camera.

let $p = (x, y, z)$

$p' = (x', y', z')$

we have

$$p'^T E p = 0$$

must be satisfied.

$\therefore$ we have

$$[x' \; y' \; z'] \; \rho \begin{bmatrix} 0 & 0 & -\sin\theta/2 \\ 0 & 0 & \cos\theta/2 \\ -\sin\theta/2 & +\cos\theta/2 & 0 \end{bmatrix}_{3\times3} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = 0$$

$$\therefore \boxed{\left(\sin\left(\frac{\theta}{2}\right)\right)(x'z + z'x) + \cos\left(\theta/2\right)\cdot(y'z - z'y) = 0}$$

Note that now, we only have $\underline{\underline{1}}$ parameter 'θ' which can be recovered as

$$\theta = -2\tan^{-1}\left(\frac{y'z - z'y}{-(x'z + z'x)}\right)$$

Thus, we only have $\underline{\underline{\text{1 parameter}}}$ and only need $\underline{\underline{\text{1 point}}}$ for estimating the motion.

Thus, $\underline{\text{1-point RANSAC}}$ works well here for estimating only '$\underline{\underline{\theta}}$'.

However we may ~~need~~ need 2 points for estimating need 2 points for estimating 'p' also. Thus 2-point will definitely walk.