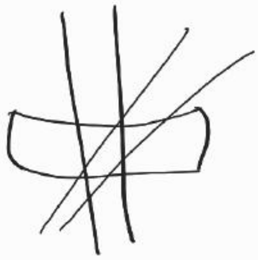
- Generalized cameras

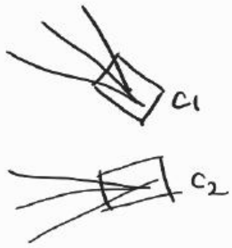
Pinhole camera : light rays converge at camera center



Generalized camera : light rays do not meet at single point
e.g. fun house mirror

- Applications

Multi-camera systems w/ minimal or w/o overlapping FoV



Although C_1, C_2 are pinhole cameras, they together are treated as generalized camera ∵ rays ✗ meet at single point ($C_1 \neq C_2$)

Advantages
- Low-cost, easy to maintain cameras (vs. LiDAR)
- Large FoV
- Can calculate absolute scale by generalized epipolar geometry

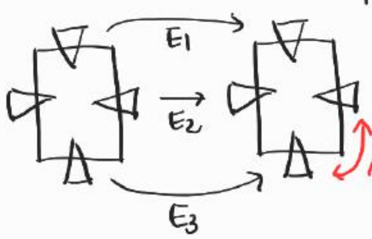
Epipolar geometry

$E \rightarrow R, t$ (5 dof)
up-to-scale

Generalized Epipolar geometry

GEM $\rightarrow R, t$ (6 dof)
absolute scale

Disadvantages
- Cannot apply stereo camera theory ∵ FoV w/o overlapping
- Processing camera independently is inefficient



Need to use relationships within cameras

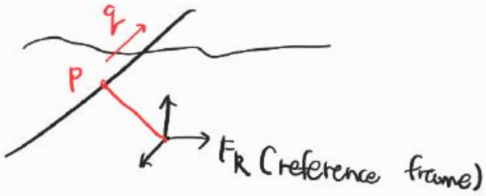
(E_1, E_2, E_3) : treated independently

- Plucker vectors

mechanisms to describe points/arbitrary lines in space/transformations of generalized systems

6D Vector
- direction $q \in R^3$: any length on a line
- moment $q' \in R^3$: $q' = q \times p$ where p is any point on a line

Constraints
- $q^T q' = 0$
- homogenous, scale is trivial



* Reference frame could be at
- any point of rigid body (car, drone)
- Center of one (of many) camera

$L = [q^T, q'^T]^T$ → often q is defined as unit vector

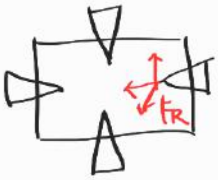
Set of all points on L are expressed as $(q \times q') + \alpha q \quad \forall \alpha \in R$

αq → signed distance

$(q \times q')$ → point on L closest to origin of F_R

- Multi-camera system

Assume reference frame is at camera center C_i



For C_i :

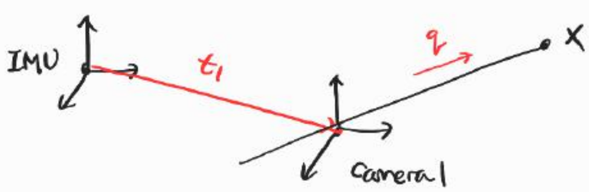
pixel (x, y) → ray $[q^T, q'^T] = [K_{C_i}^{-1}[x, y, 1]^T, 0^T]$

$q' = 0 \quad \because q \times p = q \times [0, 0, 0]^T = 0^T$

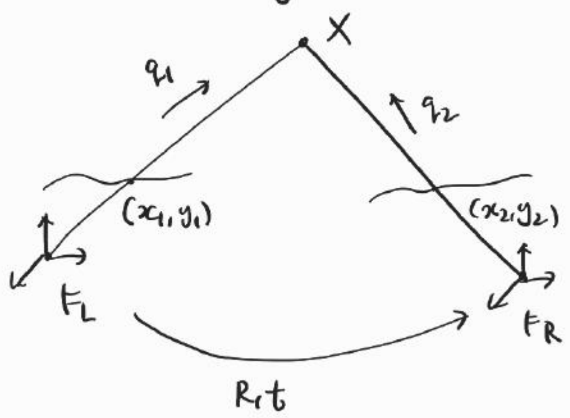
For C_j $(i \neq j)$:

pixel (x, y) → ray $[q^T, q'^T] = [R_{C_i} K_{C_i}^{-1} [x, y, 1]^T, (q \times t_{C_i})^T]$

* In practice, IMU (high freq. sensor) is used for reference frame



- Two view geometry



Point correspondence $(x_1,y_1) \leftrightarrow (x_2,y_2)$

$\parallel$

Plucker lines $[q_1^T, q_1'^T], [q_2^T, q_2'^T]^T$ intersect at X

* $\left[\begin{array}{l} R_{c_i}, t_{c_i} \text{ before : reference frame} \rightarrow \text{camera } C_i \\ R, t \text{ here : Camera left} \rightarrow \text{Camera right} \end{array}\right.$

Then, left ray can be expressed w/ right coordinate system (F_R) :

$$\begin{bmatrix} R & 0 \\ [t]_\times R & R \end{bmatrix} \begin{bmatrix} q_1 \\ q_1' \end{bmatrix} = \begin{bmatrix} Rq_1 \\ [t]_\times R q_1 + R q_1' \end{bmatrix} \quad - \text{Eq(1)}$$

Also plucker vectors a, b (w/ their own coordinate systems) intersect iff

$$\begin{bmatrix} q_b \\ q_b' \end{bmatrix}^T \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} q_a \\ q_a' \end{bmatrix} = q_b^T q_a' + q_b'^T q_a = 0 \quad - \text{Eq(2)}$$

Combining Eq(1) and Eq(2) $(q_a \leftarrow Rq_1,\ q_a' \leftarrow [t]_\times R q_1 + R q_1')$:

$$q_2^T [t]_\times R q_1 + q_2^T R q_1' + q_2'^T R q_1 = 0$$

OR

$$\begin{bmatrix} q_2 \\ q_2' \end{bmatrix}^T \begin{bmatrix} E & R \\ R & 0 \end{bmatrix} \begin{bmatrix} q_1 \\ q_1' \end{bmatrix} = 0$$

* Similar to Epipolar Geometry

$x'^T E x = 0$

$E \in \mathbb{R}^{3\times3}$

Generalized Epipolar Geometry

$\begin{bmatrix} E & R \\ R & 0 \end{bmatrix} \in \mathbb{R}^{6\times6}$ is generalized essential matrix

$[q_1^T, q_1'^T], [q_2^T, q_2'^T]^T$ are point correspondence represented by plucker coordinates

- Solving Generalized Epipolar Geometry

Solve $\begin{bmatrix} E & R \\ R & 0 \end{bmatrix}$: 18 unique entries $E(9), R(9)$ (w/o considering constraints)

$\downarrow$ they are linear to plucker coordinates

$a^T g = 0$ where a is known (q_1, q_1', q_2, q_2')

$g \in \mathbb{R}^{18\times1}$ is unknown (E, R)

$\therefore n \geq 17$ point correspondences required to solve $Ag = 0$ where $A \in \mathbb{R}^{n\times18}$

g = last column of V for $SVD(A) = U\Sigma V^T$

$= \lambda v$ $\because$ scale ambiguity

↓ Unlike epipolar geometry, scale λ can be calculated

$g = [E|R]$, since $\det(R) = 1 \rightarrow \det([\lambda v_{10}, \lambda v_{11} \cdots \lambda v_{18}]) = 0$

After solving λ, absolute scale t can be derived using $E = [t]_\times R$

∴ No scale ambiguity

- Generalized point reconstruction

3D point X intersect at $L_L = L_R$:

$\underline{R(q_1 \times q_1') + \alpha_1 R q_1 + t} = (q_2 \times q_2') + \alpha_2 q_2 \rightarrow$ Solve α_1, α_2 : Over-determined system $A[\alpha_1, \alpha_2]^T = b$

↳ Left line expressed w.r.t. right coordinate frame

$\therefore X = (q_1 \times q_1') + \alpha_1 q_1 = (q_2 \times q_2') + \alpha_2 q_2$

(F_L frame) (F_R frame)

- Analysis of Degeneracies

3 cases where degeneracy of generalized epipolar geometry occurs:

- Locally central projection
- Axial cameras
- Locally-central-and-axial cameras

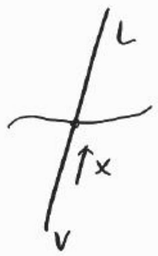


Image ray passing camera center v w/ unit direction x:

$$L = (x^T, (v \times x)^T)^T$$

↓ Then, generalized epipolar geometry (GEC) is

$$x_i^T E x_i' + x_i^T R (v_i' \times x_i') + (v_i \times x_i)^T R x_i' = 0$$

What happens for this equation for different camera geometries?

w/ sufficiently enough point correspondences, convert into linear form $Ae = 0$

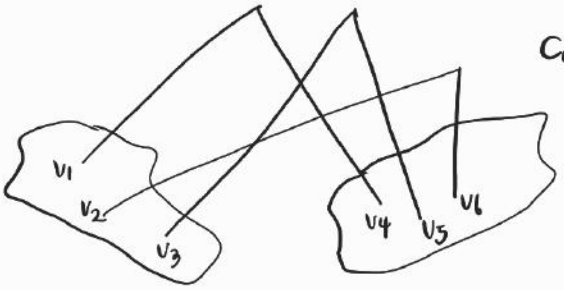
For degenerate case, $\text{rank}(A) < 17$ (18-point algo.)

↳ Suppose there are r linearly independent equations from GEC

Then $\text{rank}(A) \le r$

- General case

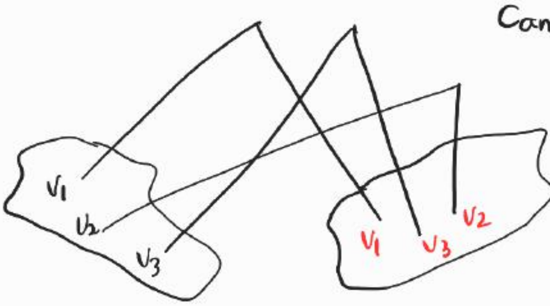
↳ Unconstrainted image rays $\equiv$ rank(A) is 17 $\equiv$ solution is unique



Camera at first view ≠ camera at second view

e.g. $v_1 \neq v_4$, $v_2 \neq v_6$

(1) Locally central projection



Camera at first view = camera at second view

∴ Same camera is tracked into different view

Correspondence $(x_i, v_i) \leftrightarrow (x_i', v_i)$ i.e. $v_i = v_i'$

$\therefore x_i^T E x_i' + x_i^T R (v_i \times x_i') + (v_i \times x_i)^T R x_i' = 0$

Solution of GEC $\begin{cases} \text{if } E \neq 0, \ (E, R) \\ \text{if } E = 0, \ (0, I) \quad \because \underline{x_i^T (v_i \times x_i') + (v_i \times x_i)^T x_i'} = 0 \end{cases}$

anti-symmetry of triple product is always 0

degenerate solution (rank(A) < 17)

Generally, the solution can be expressed in 2d linear family:

$\lambda(E, R) + \mu(0, I) = (\lambda E, \lambda R + \mu I)$

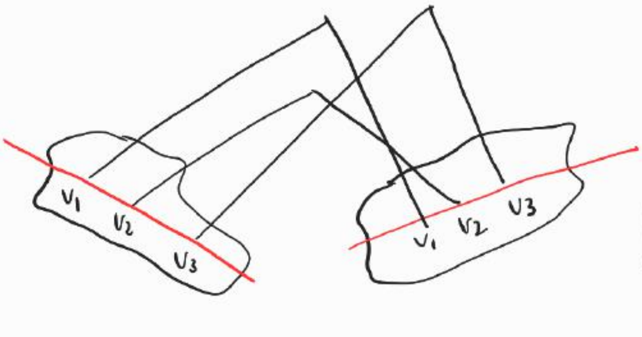
ambiguity constrained entirely on R

determined uniquely up-to-scale

* For pure translation ($R = I$), GEC degenerates into single camera epipolar geometry

$x_i^T \lambda E x_i' = 0$ ∴ cannot recover scale

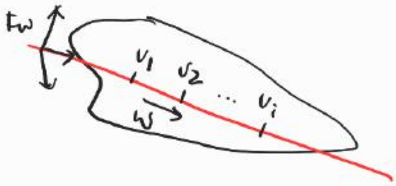
(2) Axial cameras



Occurs when cameras are aligned to single line

In practice,

- Central projection cameras
  - are rigidly mounted (as pair)
  - w/ non-overlapping FoV
  - collinear camera centers
- non-central catadioptric/fisheye cameras collinear



Assume origin of world coordinate lies on axis of camera center

If ω is direction vector of axis,

$$v_i = \alpha_i \omega \ , \ v_i' = \alpha_i' \omega$$

↓ GEC form becomes

$$x_i^T E x_i' + \alpha_i (\omega \times x_i)^T R x_i' + \alpha_i' x_i^T R (\omega \times x_i') = 0$$

This equation has rank(A) = 16 & generic solution $(\lambda E, \lambda R + \mu \omega \omega^T)$

Ambiguity only lies on R, not E ∵ World coordinate system is on axis

(3) Local - Central - and - Axial Cameras

$v_i = v_i' \ (\alpha_i = \alpha_i')$

GEC: $x_i^T E x_i' + \alpha_i (\omega \times x_i)^T R x_i' + \alpha_i x_i^T R (\omega \times x_i') = 0$ → Extra solution $(0, [\omega]_\times)$

∴ Complete solution set : $(E, R), (0, I), (0, [\omega]_\times), (0, \omega\omega^T)$

$$\rightarrow (\alpha E, \alpha R + \beta I + \gamma [\omega]_\times + \delta \omega \omega^T)$$

* This solution & non-ambiguity of E is true iff Fw origin is on the axis

- Linear algorithm under degeneracy

$$x_i^T E x_i' + x_i^T R (v_i' \times x_i') + (v_i \times x_i)^T R x_i' = 0$$

↓ w/ many point correspondences (x_i, x_i')

$Ax = A \begin{pmatrix} E \\ R \end{pmatrix} = 0$
- if rank(A) = 17, $SVD(A) = U \Sigma V^T$ / x = last col of V
- if rank(A) < 17, degeneracy occurs
  ∴ SVD approach gives wrong solutions

Approach : E remains unchanged under ambiguity

$\underset{x}{\operatorname{argmin}} \|Ax\|$ s.t. $\cancel{\|x\|=1}$ $\|E\| = 1$

$$A \begin{pmatrix} vec(E) \\ vec(R) \end{pmatrix} = (A_E \ A_R) \begin{pmatrix} vec(E) \\ vec(R) \end{pmatrix} = 0 \rightarrow A_E \, vec(E)^T + A_R \, vec(R)^T = 0$$

$$vec(R)^T = -A_R^+ A_E \, vec(E)^T$$

∴ $(A_R A_R^+ - I) A_E vec(E) = 0$ → Solve vec(E) s.t. $\|vec(E)\| = 1$

↑ Homogenous linear equation : $A' vec(E) = 0$ → $SVD(A') = U \Sigma V^T$ → unique solution E

Then, decompose E into R, R'

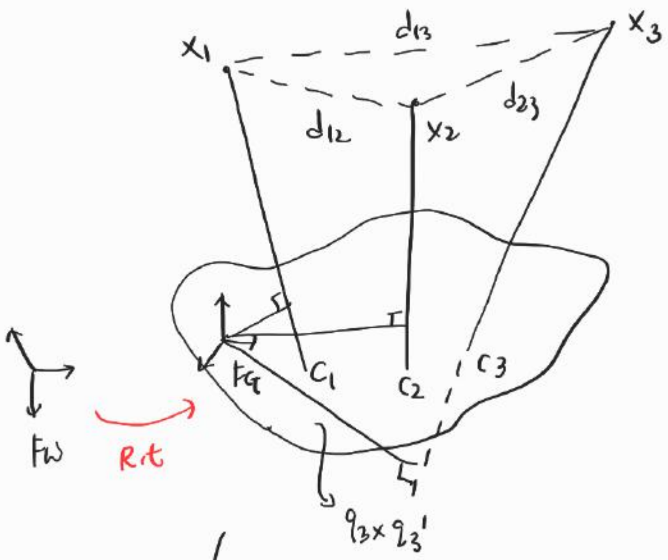
* t, t' not used here b/c of scale ambiguity.

Instead, find t, t' w/ absolute scale from GEC equation

$$x_i^T E x_i' + x_i^T R(v_i' \times x_i') + (v_i \times x_i)^T R x_i' = 0 \rightarrow A'' t = 0$$

(R, t), (R', t') : one solution only satisfy that 3D points lie front of both cameras

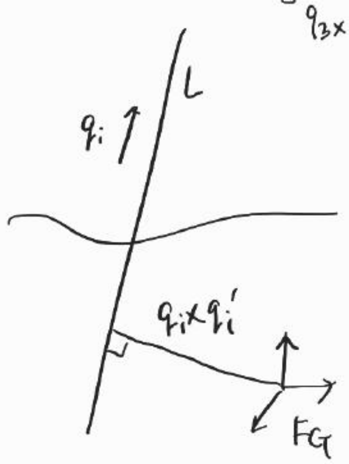
- Generalized Pose Estimation problem



Given 3D points X_i in F_W and 2D image points x_i

↓

Calculate R, t $(F_W \rightarrow F_G)$



Any point on line L r.t. F_G:

$$X_i^G = q_i \times q_i' + \lambda_i q_i$$

$\lambda_i > 0$ for 3D point to appear in front of camera

↳ signed distance b/w $q_i \times q_i'$ and X_i^G

∴ Distance b/w two 3D points d_{ij} should be equal r.t. F_G and F_W:

$$\|X_i^W - X_j^W\|^2 = \|X_i^G - X_j^G\|^2$$

$$\|X_i^W - X_j^W\|^2 = \|(q_i \times q_i' + \lambda_i q_i) - (q_j \times q_j' + \lambda_j q_j)\|^2$$

where λ are only unknowns

w/ three 3D points → 3 equations $({}_3C_2 = 3)$, 3 unknowns $(\lambda_1, \lambda_2, \lambda_3)$

Arranging equations give eight-degree polynomial of λ_3

$$A\lambda_3^8 + B\lambda_3^7 + \cdots + H\lambda_3 + I = 0$$

$\downarrow$ Using eigenvalues of companion matrix, λ_3 has 8 solutions

λ_2, λ_1 have 2 solutions each, solved using back-substitution

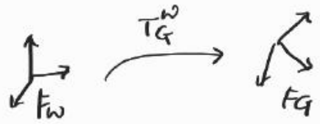
$\therefore$ $(\lambda_1, \lambda_2, \lambda_3)$ triplets have $2 \times 2 \times 8 = 32$ solutions

- Final step : selecting best-fit triplet $(\lambda_1, \lambda_2, \lambda_3)$
  - Discard solution if any of λ_i is negative / imaginary
  - For remaining triplets :
    (i) Calculate R, t $(F_W \to F_G)$ using absolute orientation (3D-3D P3P)
    (ii) Select (R_i, t_i) that gives higher inlier count (reprojection error < threshold)

- Generalized pose estimation from line correspondences

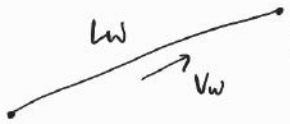
Given 2D-3D line correspondences $L_j^w \leftrightarrow l_j^c$ r.t. F_W, F_C

$\rightarrow$ Find relative transformation matrix $T_G^w = \begin{pmatrix} R_G^w & t_G^w \\ 0_{1\times3} & 1 \end{pmatrix}$



- Plucker representation of 3D Lines

↳ Use two homogenous points on lines



$P_b^w = [P_{bx}^w, P_{by}^w, P_{bz}^w, 1]^T$

$P_a^w = [P_{ax}^w, P_{ay}^w, P_{az}^w, 1]^T$

$\rightarrow$ Then, $L_W = [U_W^T, V_W^T]^T$ where

Known value

V_W : unit direction vector

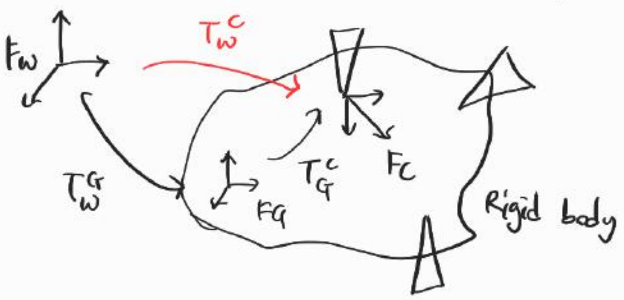
$$= \frac{P_b^w - P_a^w}{\|P_b^w - P_a^w\|}$$

$$U_W = P_a^w \times V_W$$

- Decomposing transformation matrix

$$L_C = \underbrace{T_W^C}_{\text{Unknown}} L_W = \begin{pmatrix} R_W^C & [t_W^C]_\times R_W^C \\ 0_{3\times3} & R_W^C \end{pmatrix} L_W$$

$$T_W^C = T_G^C T_W^G = \begin{pmatrix} R_G^C & t_G^C \\ 0_{1\times3} & 1 \end{pmatrix}\begin{pmatrix} R_W^G & t_W^G \\ 0_{1\times3} & 1 \end{pmatrix} = \begin{pmatrix} R_G^C R_C^W & R_G^C t_W^C + t_G^C \\ 0_{1\times3} & 1 \end{pmatrix}$$

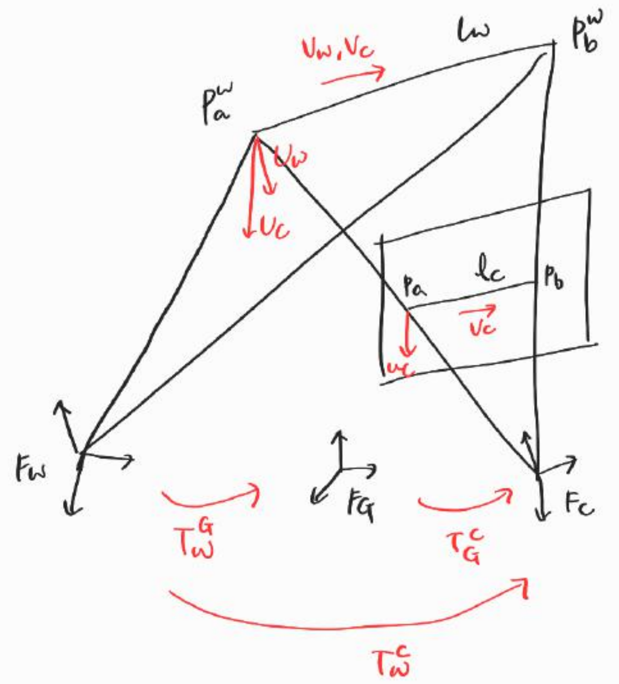


Since $L_c = [U_c^T\ V_c^T]^T \in \mathbb{R}^{6\times1}$, $U_c = \begin{pmatrix} R_w^c & [t_w^c]_\times R_w^c \\ \cancel{0_{3\times3}} & \cancel{R_w^c} \end{pmatrix} L_w = (R_w^c\ \ [t_w^c]_\times R_w^c)\begin{pmatrix} U_w \\ V_w \end{pmatrix}$

Further properties : $U_c \cdot V_c = U_c^T V_c = 0$

$$V_c = \begin{pmatrix} \cancel{R_w^c} & \cancel{[t_w^c]_\times R_w^c} \\ 0_{3\times3} & R_w^c \end{pmatrix} L_w = R_w^c V_w \rightarrow U_c R_w^c V_w = 0$$

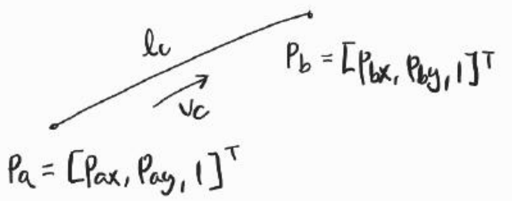
• Plucker representation of 2D lines



Known : T_G^c (Extrinsics of camera)
$L_w = [U_w^T, V_w^T]^T$
L_c, l_c (shown below)

$\downarrow$

Unknown : T_w^G, T_w^c



$\rightarrow$ Then, $l_c = [u_c^T, v_c^T]^T$ where

$$v_c = \frac{\hat{P}_b - \hat{P}_a}{\|\hat{P}_b - \hat{P}_a\|}, \quad u_c = \hat{P}_a \times v_c$$

$(\hat{P}_a = K^{-1} P_a,\ \hat{P}_b = K^{-1} P_b)$

Since $U_c^T R_w^c V_w = 0$ and $R_w^c = R_G^c R_w^G$, $U_c // u_c$:

$u_c^T R_G^c \underbrace{R_w^G}_{\text{Unknown}} V_w = 0$

• Solving R_G^w — homogenous linear equation

$ar = 0$

↳ $\mathbb{R}^{9\times1}$ of unknown R_G^w

↳ $\mathbb{R}^{1\times9}$ of known u_c^T, R_G^c, V_w

$\rightarrow$ w/ $n \geq 8$ 2D–3D line correspondences

$Ar = 0$ $(A \in \mathbb{R}^{n\times9})$

$SVD(A) = U\Sigma V^T \rightarrow r =$ last column of V

• Solving t_G^w

Since $u_c // U_c$, $\lambda u_c = U_c$

$$u_c = \left(R_w^c \quad \underbrace{[t_w^c]_\times}_{\text{Unknown}} R_w^c \right) \begin{pmatrix} U_w \\ V_w \end{pmatrix} = \underbrace{\lambda}_{\text{Unknown}} u_c$$

$\downarrow$ cross product of u_c to remove λ

$$[u_c]_\times \left(R_w^c \quad [t_w^c]_\times R_w^c \right) \begin{pmatrix} U_w \\ V_w \end{pmatrix} = 0$$

$\downarrow$ homogenous linear equation

$bt = 0$ where $t = [t_x, t_y, t_z, 1]^T$

w/ $n \geq 8$ 2D-3D line correspondences (can reuse from solving $Ar = 0$):

$Bt = 0$ (overdetermined system: $B \in \mathbb{R}^{n\times 4}$, $t \in \mathbb{R}^{4\times 1}$)

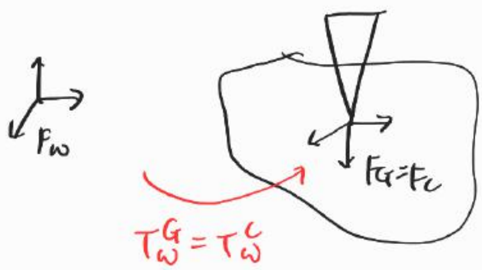
$SVD(B) = U\Sigma V^T \rightarrow t = \alpha$(last column of V)

$\uparrow$

ambiguity can be solved : $\alpha V_{i4} = 1$ (last entry of t is 1)

- Special cases

1) One camera



Camera extrinsics (R_G^c, t_G^c) vanishes

$\therefore$ Directly solve R_w^c w/o decomposition

i.e. $u_c' R_G^c R_w^G V_w \rightarrow u_c' R_w^c V_w = 0$

2) parallel 3D lines

All 3D lines have equal V_w $\rightarrow$ degeneracy where $\text{rank}(A) < 8$ for $Ar = 0$

$\therefore R_w^G$ cannot be solved

$\downarrow$ Prevention

Omit parallel 3D lines