Fundamental and Essential Matrix:

Fundamental → For Uncalib cam

Essential → For calib cam

(*) Both give info about relative orientation b/w 2 images → From pt. correspondences alone

→ Are homogeneous 3x3 matrices

(*) Rank deficiency in both matrices to model for the coplanarity constraint ⇒ $\mathrm{Rank}(E) = 2$, $\mathrm{Rank}(F) = 2$

All points on img ⇓ All lie on a plane

Must hold for all pt. correspondences

$x_{i1}^T F x_{i2} = 0$ — i^{th} corresponding points

x_{i1} → In img 1; x_{i2} → In img 2 → Points in pixel coordinate system

$x_{i1}^T E x_{i2} = 0$ → Points in camera coordinate system

(*) Coplanarity constraints across all pts. gives us: $\boxed{Af = 0}$ → Soln. from SVD

(*) 8-point algo for F (Uncalibrated)

(*) 5-point algo for E (Calibrated)

→ Only 5 pts. needed

→ Decomposed to give $\boxed{B}$ and $\boxed{R}$

- B → Baseline vector (Only direction not lengths) → Along which direction is cam 2 wrt 1 but not how far
- R → Orientation of cam 2 wrt 1

(*) If E/F are known, the search for pt. correspondences reduces to a single line in the other img called epipolar line

(*) E has 5 DoF → 3 Rotation + 2 Translation → Depth missing

(*) F has 7 DoF → 3 Rotation + 2 Translation + 2 Calib params

⊛ For a pair of calibrated cameras [Angle-preserving mapping]
- We need to estimate R (3) and X_0 (3) for each camera, i.e., 6 per camera
- Hence, involves solving for $\boxed{\text{12 params} \rightarrow \text{Calibrated}}$

⊛ For a pair of uncalibrated camera [Straight-line preserving mapping]
- We additionally need to find 5 extra params (intrinsics) for each camera
- 10 additional for both cameras combined
- Hence, involves solving for $\boxed{\text{22 params} \rightarrow \text{Uncalibrated}}$

⊛ Single cam calib needs knowledge about scene
- P3P, DLT need control point locations in the scene

From Two Cameras:

⊛ $\boxed{\text{No scale info}}$ without scene knowledge

⊛ We only get the relative transform b/w the 2 cams, that too without scale

⊛ We also $\boxed{\text{do not get absolute configuration of cam 1 wrt scene}}$

⊛ For calibrated cams, we lose $\underbrace{1}_{\text{Scale}} + \underbrace{\overbrace{6}}_{R, X_0 \text{ globally}} = 7$ directly and no way to estimate

⊛ Hence, 5 params (12-7) are only solvable in E → $\boxed{\text{E has 5 DoF}}$

⊛ This is a similarity transform - Angle-preserving transform

⊛ $\boxed{\text{3D model of the scene known upto a similarity transform}}$
↳ Photogrammetric model

⊛ Upto this, we have the $\boxed{\text{relative orientation problem}}$

⊛ $\boxed{\text{Fixing the scale transform and where cam 1 is in the world is the absolute orientation problem}}$ → Requires knowledge of 3 3D points

* To build the 3D scene model, we will need one of these:
  - Location of three 3D points in the scene
  - Location of projection centers of both cams in the world
  - Length of baseline vector gives location of cam 2 (6 more) and the last param is the length (or) orientation — Can get from triangulation estimates
  - Distances in 3D world b/w points in the real world

Uncalibrated cams:

* Straight-line preserving and not angle-preserving
* Reconstruction can be only done till a projective transform

$$\underset{4\times4}{P} = \begin{bmatrix} R & t \\ a^T & 1 \end{bmatrix}_{4\times4}$$

a^T → Handles the projection

To know this, we need 15 params, 1 param scaled to 1 for homogeneity in the 4×4 P matrix

* We are left with 22-15=7 params that can be estimated

$\boxed{\text{F has 7 DoF}}$

* Hence, to build the 3D model, we need 5 control points in the scene to get the 15 params (5×3 coords)
* Scalar triple product to model the coplanarity constraint
* Normalized ~~coord~~ world coordinate → Direction of world coordinate in 3D camera frame
  - Origin translated to projective center
* For $b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$, $S_b = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}$ from cross-product math

  S_b is skew-symmetric
* Final coplanarity constraint $\boxed{{}^n x'^T S_b {}^n x'' = 0}$ → ${}^n x$ → Normalised world coord
  - Only gives direction.

$${}^{n}x' = (R')^{-1}(K')^{-1}x'$$

$${}^{n}x'^{T} S_b\, {}^{n}x'' = 0$$

$$\Rightarrow x'^{T}(K')^{-T}(R')^{-T} S_b (R'')^{-1}(K'')^{-1} x'' = 0$$

$$\Rightarrow \boxed{F = (K')^{-T} R' S_b R''^{T}(K'')^{-1}}$$

Hence collinearity constraint is:

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

Calibrated Cams:

$$\boxed{E = R' S_b R''^{T}}$$

$$\boxed{{}^{k}x'^{T} E\, {}^{k}x'' = 0}$$

Direction vector of the 3D point in the cam. frame

$$\boxed{{}^{k}x' = (K')^{-1} x'}$$

Pixel coordinates of the corresponding point in the 2D image frame

We have the 5 params from 12−7=5, 7 which we cannot obtain without scene knowledge as it is a similarity transform — 3 Rot + 3 Trans + 1 Scale

$$S = \begin{bmatrix} mR & t \\ 0^{T} & 1 \end{bmatrix}$$

In the 3x3 E matrix, we already have 5 params, and hence we have

$\boxed{9-5=4 \text{ constraints to ensure } E \text{ is homogeneous and singular}}$