3D Computer Vision

Radim Šára Martin Matoušek

Center for Machine Perception
Department of Cybernetics
Faculty of Electrical Engineering
Czech Technical University in Prague

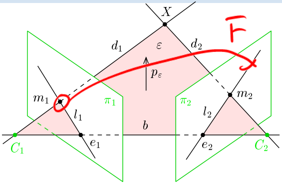
https://cw.fel.cvut.cz/wiki/courses/tdv/start

http://cmp.felk.cvut.cz
mailto:sara@cmp.felk.cvut.cz
phone ext. 7203

rev. November 2, 2021

Open Informatics Master's Course

►Expressing Epipolar Constraint Algebraically



$$\mathbf{P}_i = \begin{bmatrix}\mathbf{Q}_i & \mathbf{q}_i\end{bmatrix} = \mathbf{K}_i \begin{bmatrix}\mathbf{R}_i & \mathbf{t}_i\end{bmatrix}, \ i = 1, 2$$

$\mathbf{R}_{21}$ – relative camera rotation, $\mathbf{R}_{21} = \mathbf{R}_2 \mathbf{R}_1^\top$

$\mathbf{t}_{21}$ – relative camera translation, $\mathbf{t}_{21} = \mathbf{t}_2 - \mathbf{R}_{21}\mathbf{t}_1 = -\mathbf{R}_2 \mathbf{b}$ →74

$\mathbf{b}$ – baseline vector (world coordinate system)

remember: $\mathbf{C} = -\mathbf{Q}^{-1}\mathbf{q} = -\mathbf{R}^\top \mathbf{t}$ →33 and 35

$$\boxed{0 = \mathbf{d}_2^\top \underbrace{\mathbf{p}_\varepsilon}_{\text{normal of } \varepsilon} \simeq \underbrace{(\mathbf{Q}_2^{-1}\underline{\mathbf{m}}_2)^\top}_{\text{optical ray}} \underbrace{\mathbf{Q}_1^\top \underline{\mathbf{l}}_1}_{\text{optical plane}} = \underline{\mathbf{m}}_2^\top \underbrace{\mathbf{Q}_2^{-\top}\mathbf{Q}_1^\top(\underline{\mathbf{e}}_1 \times \underline{\mathbf{m}}_1)}_{\text{image of } \varepsilon \text{ in } \pi_2} = \underline{\mathbf{m}}_2^\top \underbrace{(\mathbf{Q}_2^{-\top}\mathbf{Q}_1^\top [\underline{\mathbf{e}}_1]_\times)}_{\text{fundamental matrix } \mathbf{F}} \underline{\mathbf{m}}_1}$$

Epipolar constraint $\quad \underline{\mathbf{m}}_2^\top \mathbf{F}\, \underline{\mathbf{m}}_1 = 0 \quad$ is a point-line incidence constraint

- point $\underline{\mathbf{m}}_2$ is incident on epipolar line $\underline{\mathbf{l}}_2 \simeq \mathbf{F}\underline{\mathbf{m}}_1$
- point $\underline{\mathbf{m}}_1$ is incident on epipolar line $\underline{\mathbf{l}}_1 \simeq \mathbf{F}^\top\underline{\mathbf{m}}_2$
- $\mathbf{F}\underline{\mathbf{e}}_1 = \mathbf{F}^\top\underline{\mathbf{e}}_2 = \mathbf{0}$ (non-trivially)
- all epipolars meet at the epipole

$$\underline{\mathbf{e}}_1 \simeq \mathbf{Q}_1\mathbf{C}_2 + \mathbf{q}_1 = \mathbf{Q}_1\mathbf{C}_2 - \mathbf{Q}_1\mathbf{C}_1 = \mathbf{K}_1\mathbf{R}_1\mathbf{b} = -\mathbf{K}_1\mathbf{R}_1\mathbf{R}_2^\top\mathbf{t}_{21} = \cancel{-}\mathbf{K}_1\mathbf{R}_{21}^\top\mathbf{t}_{21}$$

E

$$\mathbf{F} = \mathbf{Q}_2^{-\top}\mathbf{Q}_1^\top[\underline{\mathbf{e}}_1]_\times = \mathbf{Q}_2^{-\top}\mathbf{Q}_1^\top[-\mathbf{K}_1\mathbf{R}_{21}^\top\mathbf{t}_{21}]_\times = \overset{\circledast 1}{\cdots} \simeq \mathbf{K}_2^{-\top}[-\mathbf{t}_{21}]_\times \mathbf{R}_{21}\mathbf{K}_1^{-1} \quad \text{fundamental}$$

$$\mathbf{E} = [-\mathbf{t}_{21}]_\times \mathbf{R}_{21} = \underbrace{[\mathbf{R}_2\mathbf{b}]_\times}_{\text{baseline in Cam 2}} \mathbf{R}_{21} = \mathbf{R}_{21} \underbrace{[\mathbf{R}_1\mathbf{b}]_\times}_{\text{baseline in Cam 1}} = \mathbf{R}_{21}[-\mathbf{R}_{21}^\top\mathbf{t}_{21}]_\times \quad \text{essential}$$

►The Structure and the Key Properties of the Fundamental Matrix

$0 = \underline{m}_2^\top \mathbf{F} \underline{m}_1 = \underline{m}_2^\top \mathbf{H}_e^{-\top} (e_1 \times \underline{m}_1)$

$$\mathbf{F} = (\underbrace{\mathbf{Q}_2\mathbf{Q}_1^{-1}}_{\text{epipolar homography } \mathbf{H}_e})^{-\top}[\mathbf{e}_1]_\times = \underbrace{\mathbf{K}_2^{-\top}\mathbf{R}_{21}\mathbf{K}_1^{\top}}_{\mathbf{H}_e^{-\top}}\overbrace{[\mathbf{e}_1]_\times}^{\text{left epipole}} \overset{\to 76}{\simeq} \overbrace{[\mathbf{H}_e\mathbf{e}_1]_\times}^{\text{right epipole}}\mathbf{H}_e = \mathbf{K}_2^{-\top}\underbrace{[-\mathbf{t}_{21}]_\times\mathbf{R}_{21}}_{\text{essential matrix } \mathbf{E}}\mathbf{K}_1^{-1}$$

1. $\mathbf{E}$ captures relative camera pose only [Longuet-Higgins 1981]

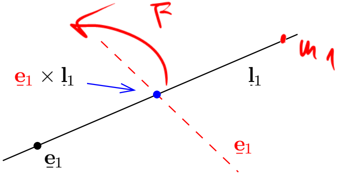
   (the change of the world coordinate system does not change $\mathbf{E}$)

$$\begin{bmatrix}\mathbf{R}'_i & \mathbf{t}'_i\end{bmatrix} = \begin{bmatrix}\mathbf{R}_i & \mathbf{t}_i\end{bmatrix} \cdot \begin{bmatrix}\mathbf{R} & \mathbf{t} \\ \mathbf{0}^\top & 1\end{bmatrix} = \begin{bmatrix}\mathbf{R}_i\mathbf{R} & \mathbf{R}_i\mathbf{t} + \mathbf{t}_i\end{bmatrix},$$

then

$$\mathbf{R}'_{21} = \mathbf{R}'_2\mathbf{R}'^{\top}_1 = \cdots = \mathbf{R}_{21} \qquad \mathbf{t}'_{21} = \mathbf{t}'_2 - \mathbf{R}'_{21}\mathbf{t}'_1 = \cdots = \mathbf{t}_{21}$$

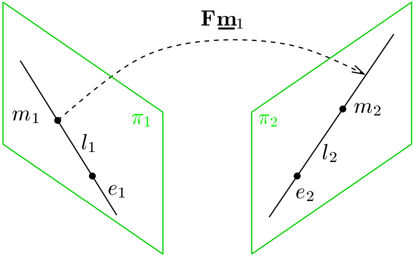
2. the translation length $\mathbf{t}_{21}$ is <u>lost</u> since $\mathbf{E}$ is homogeneous
3. $\mathbf{F}$ maps points to lines and it is not a homography
4. $\mathbf{H}_e$ maps epipoles to epipoles, $\mathbf{H}_e^{-\top}$ epipolar lines to epipolar lines: $\underline{\mathbf{l}}_2 \simeq \mathbf{H}_e^{-\top}\underline{\mathbf{l}}_1$



another epipolar line map: $\underline{\mathbf{l}}_2 \simeq \mathbf{F}[\underline{\mathbf{e}}_1]_\times\underline{\mathbf{l}}_1$

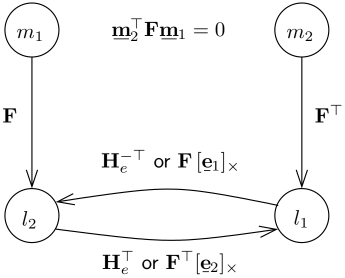
- proof by point/line 'transmutation' (left)
- point $\underline{\mathbf{e}}_1$ does not lie on line $\underline{\mathbf{e}}_1$ (dashed): $\underline{\mathbf{e}}_1^\top\underline{\mathbf{e}}_1 \neq 0$
- $\mathbf{F}[\underline{\mathbf{e}}_1]_\times$ is not a homography, unlike $\mathbf{H}_e^{-\top}$ but it does the same job for epipolar line mapping
- no need to decompose $\mathbf{F}$ to obtain $\mathbf{H}_e$

►Summary: Relations and Mappings Involving Fundamental Matrix



$$0 = \underline{\mathbf{m}}_2^\top \mathbf{F}\, \underline{\mathbf{m}}_1$$

$$
\begin{aligned}
\underline{\mathbf{e}}_1 &\simeq \mathrm{null}(\mathbf{F}), & \underline{\mathbf{e}}_2 &\simeq \mathrm{null}(\mathbf{F}^\top)\\
\underline{\mathbf{e}}_1 &\simeq \mathbf{H}_e^{-1}\underline{\mathbf{e}}_2 & \underline{\mathbf{e}}_2 &\simeq \mathbf{H}_e\underline{\mathbf{e}}_1\\
\underline{\mathbf{l}}_1 &\simeq \mathbf{F}^\top\underline{\mathbf{m}}_2 & \underline{\mathbf{l}}_2 &\simeq \mathbf{F}\underline{\mathbf{m}}_1\\
\underline{\mathbf{l}}_1 &\simeq \mathbf{H}_e^\top\underline{\mathbf{l}}_2 & \underline{\mathbf{l}}_2 &\simeq \mathbf{H}_e^{-\top}\underline{\mathbf{l}}_1\\
\underline{\mathbf{l}}_1 &\simeq \mathbf{F}^\top[\underline{\mathbf{e}}_2]_\times\underline{\mathbf{l}}_2 & \underline{\mathbf{l}}_2 &\simeq \mathbf{F}[\underline{\mathbf{e}}_1]_\times\underline{\mathbf{l}}_1
\end{aligned}
$$



- $\mathbf{F}[\underline{\mathbf{e}}_1]_\times$ maps epipolar lines to epipolar lines but it is not a homography
- $\mathbf{H}_e = \mathbf{Q}_2\mathbf{Q}_1^{-1}$ is the epipolar homography→78
  $\mathbf{H}_e^{-\top}$ maps epipolar lines to epipolar lines, where

$$\mathbf{H}_e = \mathbf{Q}_2\mathbf{Q}_1^{-1} = \mathbf{K}_2\mathbf{R}_{21}\mathbf{K}_1^{-1}$$

you have seen this →59

►Representation Theorem for Fundamental Matrices

Def: $\mathbf{F}$ is fundamental when $\mathbf{F} \simeq \mathbf{H}^{-\top}[\underline{\mathbf{e}}_1]_\times$, where $\mathbf{H}$ is regular and $\underline{\mathbf{e}}_1 \simeq \operatorname{null}\mathbf{F} \neq \mathbf{0}$.

Theorem: A 3×3 matrix $\mathbf{A}$ is fundamental iff it is of rank 2.

Proof.
Direct: By the geometry, $\mathbf{H}$ is full-rank, $\underline{\mathbf{e}}_1 \neq \mathbf{0}$, hence $\mathbf{H}^{-\top}[\underline{\mathbf{e}}_1]_\times$ is a 3×3 matrix of rank 2.
Converse:

1. let $\mathbf{A} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ be the SVD of $\mathbf{A}$ of rank 2; then $\mathbf{D} = \operatorname{diag}(\lambda_1, \lambda_2, 0)$, $\lambda_1 \geq \lambda_2 > 0$
2. we write $\mathbf{D} = \mathbf{B}\mathbf{C}$, where $\mathbf{B} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, $\mathbf{C} = \operatorname{diag}(1, 1, 0)$ $\lambda_3 \neq 0$
3. then $\boxed{\mathbf{A} = \mathbf{U}\mathbf{B}\mathbf{C}\mathbf{V}^\top} = \mathbf{U}\mathbf{B}\mathbf{C}\underbrace{\mathbf{W}\mathbf{W}^\top}_{\mathbf{I}}\mathbf{V}^\top$ with $\mathbf{W}$ rotation
4. we look for a rotation $\mathbf{W}$ that maps $\mathbf{C}$ to a skew-symmetric $\mathbf{S}$, i.e. $\mathbf{S} = \mathbf{C}\mathbf{W}$
5. then $\mathbf{W} = \begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $|\alpha| = 1$, and $\mathbf{S} = [\mathbf{s}]_\times$, $\mathbf{s} = (0, 0, 1)$
6. we write $\mathbf{v}_3$ – 3rd column of $\mathbf{V}$, $\mathbf{u}_3$ – 3rd column of $\mathbf{U}$
$$\mathbf{A} = \mathbf{U}\mathbf{B}[\mathbf{s}]_\times\mathbf{W}^\top\mathbf{V}^\top \overset{\circledast 1}{=} \cdots = \underbrace{\mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top}_{\simeq \mathbf{H}^{-\top}}[\mathbf{v}_3]_\times \simeq \underbrace{[\mathbf{H}\mathbf{v}_3]_\times}_{\simeq[\mathbf{u}_3]_\times}\mathbf{H}, \tag{12}$$
7. $\mathbf{H}$ regular, $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$, $\mathbf{u}_3^\top\mathbf{A} = \mathbf{0}^\top$ for $\mathbf{v}_3 \neq \mathbf{0}$, $\mathbf{u}_3 \neq \mathbf{0}$ □

- we also got a (non-unique: α, λ_3) decomposition formula for fundamental matrices
- it follows there is no constraint on $\mathbf{F}$ except for the rank

Theorem

Let $\mathbf{E}$ *be a* 3×3 *matrix with SVD* $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$. *Then* $\mathbf{E}$ *is essential iff* $\mathbf{D} \simeq \mathrm{diag}(1, 1, 0)$.

Proof.

$D = \lambda \, \mathrm{diag}(1, 1, 0)$

Direct:

If $\mathbf{E}$ is an essential matrix, then the epipolar homography matrix is a rotation matrix (→78), hence $\mathbf{H}^{-\top} \simeq \mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top$ in (12) must be (λ-scaled) orthogonal, therefore $\mathbf{B} = \lambda\mathbf{I}$.

we have fixed the missing λ_3 in (12)

Then

$$\boxed{\mathbf{R}_{21} = \mathbf{H}^{-\top} \simeq \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top \simeq \mathbf{U}\mathbf{W}\mathbf{V}^\top}$$

2 solutions

$|\alpha| = 1$

Converse:

$\mathbf{E}$ is fundamental with

$$\mathbf{D} = \mathrm{diag}(\lambda, \lambda, 0) = \underbrace{\lambda\mathbf{I}}_{\mathbf{B}} \underbrace{\mathrm{diag}(1, 1, 0)}_{\mathbf{D}}$$

then $\mathbf{B} = \lambda\mathbf{I}$ in (12) and $\mathbf{U}(\mathbf{V}\mathbf{W})^\top$ is orthogonal, as required. □

▶Essential Matrix Decomposition

We are decomposing $\mathbf{E}$ to $\mathbf{E} \simeq [-\mathbf{t}_{21}]_\times \mathbf{R}_{21} = \mathbf{R}_{21}[-\mathbf{R}_{21}^\top \mathbf{t}_{21}]_\times$ [H&Z, sec. 9.6]

1. compute SVD of $\mathbf{E} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$ and verify $\mathbf{D} = \lambda\,\mathrm{diag}(1,1,0)$
2. ensure $\mathbf{U}$, $\mathbf{V}$ are rotation matrices by $\mathbf{U} \mapsto \det(\mathbf{U})\mathbf{U}$, $\mathbf{V} \mapsto \det(\mathbf{V})\mathbf{V}$
3. compute

$$\mathbf{R}_{21} = \mathbf{U}\underbrace{\begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}}_{\mathbf{W}}\mathbf{V}^\top, \quad \mathbf{t}_{21} = -\beta\,\mathbf{u}_3, \qquad |\alpha| = 1, \quad \beta \neq 0 \tag{13}$$

Notes

- $\mathbf{v}_3 \simeq \mathbf{R}_{21}^\top \mathbf{t}_{21}$ by (12), hence $\mathbf{R}_{21}\mathbf{v}_3 \simeq \mathbf{t}_{21} \simeq \mathbf{u}_3$ since it must fall in left null space by $\mathbf{E} \simeq [\mathbf{u}_3]_\times \mathbf{R}_{21}$
- $\mathbf{t}_{21}$ is recoverable up to scale β and direction $\mathrm{sign}\,\beta$
- the result for $\mathbf{R}_{21}$ is unique up to $\alpha = \pm 1$ despite non-uniqueness of SVD
- the change of sign in α rotates the solution by 180° about $\mathbf{t}_{21}$

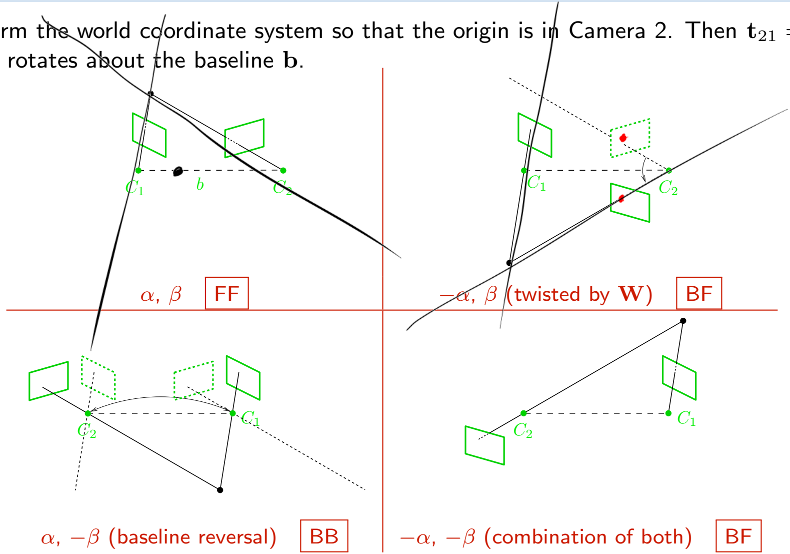
  $\mathbf{R}(\alpha) = \mathbf{U}\mathbf{W}\mathbf{V}^\top$, $\mathbf{R}(-\alpha) = \mathbf{U}\mathbf{W}^\top\mathbf{V}^\top \Rightarrow \mathbf{T} = \mathbf{R}(-\alpha)\mathbf{R}^\top(\alpha) = \cdots = \mathbf{U}\,\mathrm{diag}(-1,-1,1)\mathbf{U}^\top$
  which is a rotation by 180° about $\mathbf{u}_3 \simeq \mathbf{t}_{21}$: show that $\mathbf{u}_3$ is the rotation axis

$$\mathbf{U}\,\mathrm{diag}(-1,-1,1)\mathbf{U}^\top\mathbf{u}_3 = \mathbf{U}\begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{u}_3$$

- 4 solution sets for 4 sign combinations of α, β see next for geometric interpretation

▶Four Solutions to Essential Matrix Decomposition

Transform the world coordinate system so that the origin is in Camera 2. Then $\mathbf{t}_{21} = -\mathbf{b}$ and $\mathbf{W}$ rotates about the baseline $\mathbf{b}$. →77



- chirality constraint: all 3D points are in front of both cameras
- this singles-out the upper left case [H&Z, Sec. 9.6.3]

►7-Point Algorithm for Estimating Fundamental Matrix

Problem: Given a set $\{(x_i, y_i)\}_{i=1}^{k}$ of $k = 7$ finite correspondences, estimate f. m. $\mathbf{F}$.

$$\underline{\mathbf{y}}_i^\top \mathbf{F}\, \underline{\mathbf{x}}_i = 0, \quad i = 1, \dots, k, \qquad \underline{\text{known}}:\ \ \underline{\mathbf{x}}_i = (u_i^1, v_i^1, 1), \ \ \underline{\mathbf{y}}_i = (u_i^2, v_i^2, 1)$$

terminology: correspondence = truth, later: match = algorithm's result; hypothesized corresp.

Solution:

$$0 = \underline{\mathbf{y}}_i^\top \mathbf{F}\, \underline{\mathbf{x}}_i = (\underline{\mathbf{y}}_i \underline{\mathbf{x}}_i^\top) : \mathbf{F} = \left(\operatorname{vec}(\underline{\mathbf{y}}_i \underline{\mathbf{x}}_i^\top)\right)^\top \operatorname{vec}(\mathbf{F}),$$

rotation property of matrix trace →71

$$\operatorname{vec}(\mathbf{F}) = \begin{bmatrix} f_{11} & f_{21} & f_{31} & \dots & f_{33} \end{bmatrix}^\top \in \mathbb{R}^9$$

column vector from matrix

$$\mathbf{D} = \begin{bmatrix} \left(\operatorname{vec}(\mathbf{y}_1 \mathbf{x}_1^\top)\right)^\top \\ \left(\operatorname{vec}(\mathbf{y}_2 \mathbf{x}_2^\top)\right)^\top \\ \left(\operatorname{vec}(\mathbf{y}_3 \mathbf{x}_3^\top)\right)^\top \\ \vdots \\ \left(\operatorname{vec}(\mathbf{y}_k \mathbf{x}_k^\top)\right)^\top \end{bmatrix} = \underbrace{\begin{bmatrix} u_1^1 u_1^2 & u_1^1 v_1^2 & u_1^1 & u_1^2 v_1^1 & v_1^1 v_1^2 & v_1^1 & u_1^2 & v_1^2 & 1 \\ u_2^1 u_2^2 & u_2^1 v_2^2 & u_2^1 & u_2^2 v_2^1 & v_2^1 v_2^2 & v_2^1 & u_2^2 & v_2^2 & 1 \\ u_3^1 u_3^2 & u_3^1 v_3^2 & u_3^1 & u_3^2 v_3^1 & v_3^1 v_3^2 & v_3^1 & u_3^2 & v_3^2 & 1 \\ \vdots & & & & & & & & \vdots \\ u_k^1 u_k^2 & u_k^1 v_k^2 & u_k^1 & u_k^2 v_k^1 & v_k^1 v_k^2 & v_k^1 & u_k^2 & v_k^2 & 1 \end{bmatrix}}_{7\times 9} \in \mathbb{R}^{k,9}$$

$$\boxed{\mathbf{D}\, \underbrace{\operatorname{vec}(\mathbf{F})}_{\mathbb{R}^9} = \mathbf{0}}$$

$a = \mathrm{vec}(A)$ $\quad A = \mathrm{vec}^{-1}(a) = \mathrm{mat}(a)$ $\quad \alpha F_1 + (1-\alpha) F_2 = F$

$$\mathbf{D}\,\mathrm{vec}(\mathbf{F}) = \mathbf{0}, \quad \mathbf{D} \in \mathbb{R}^{k,9}$$

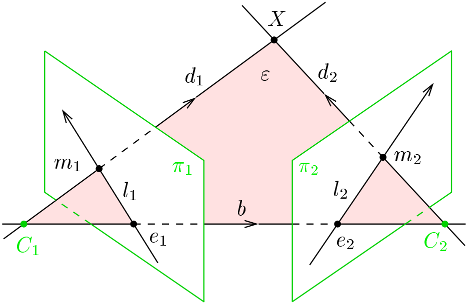
- for $k = 7$ we have a rank-deficient system, the null-space of $\mathbf{D}$ is 2-dimensional
- but we know that $\det \mathbf{F} = 0$, hence
  1. find a basis of the null space of $\mathbf{D}$: $\mathbf{F}_1$, $\mathbf{F}_2$ — by SVD or QR factorization
  2. get up to 3 real solutions for α from
     $$\det(\alpha \mathbf{F}_1 + (1-\alpha)\mathbf{F}_2) = 0$$
     cubic equation in α
  3. get up to 3 fundamental matrices $\mathbf{F}_i = \alpha_i \mathbf{F}_1 + (1-\alpha_i)\mathbf{F}_2$
  4. if $\operatorname{rank} \mathbf{F}_i < 2$ for all $i = 1, 2, 3$ then fail

$\begin{bmatrix} \alpha F_{111} + (1-\alpha) F_{211} & \cdot & \cdot \\ \alpha F_{121} + (1-\alpha) F_{221} & \cdot & \cdot \end{bmatrix}$

- the result may depend on image (domain) transformations
- normalization improves conditioning →92
- this gives a good starting point for the full algorithm →104
- dealing with mismatches need not be a part of the 7-point algorithm →105

A Note on Oriented Epipolar Constraint

- a tighter epipolar constraint preserves orientations
- requires all points and cameras be on the same side of the plane at infinity



$\underline{m}_2^\top F \underline{m}_1 = 0$

$$\boxed{(\underline{\mathbf{e}}_2 \times \underline{\mathbf{m}}_2) \overset{+}{\sim} \mathbf{F}\,\underline{\mathbf{m}}_1}$$

notation: $\underline{\mathbf{m}} \overset{+}{\sim} \underline{\mathbf{n}}$ means $\underline{\mathbf{m}} = \lambda \underline{\mathbf{n}}$, $\lambda > 0$

- we can read the constraint as $(\underline{\mathbf{e}}_2 \times \underline{\mathbf{m}}_2) \overset{+}{\sim} \mathbf{H}_e^{-\top} (\mathbf{e}_1 \times \underline{\mathbf{m}}_1)$
- note that the constraint is not invariant to the change of either sign of $\underline{\mathbf{m}}_i$
- all 7 correspondence in 7-point alg. must have the same sign — see later
- this may help reject some wrong matches, see $\rightarrow$105 — [Chum et al. 2004]
- an even more tight constraint: scene points <u>in front</u> of both cameras — expensive

  this is called <u>chirality constraint</u>

Thank You