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. October 24, 2023

Open Informatics Master's Course

Module IV

Computing with a Camera Pair

covered by

[1] [H&Z] Secs: 9.1, 9.2, 9.6, 11.1, 11.2, 11.9, 12.2, 12.3, 12.5.1

[2] H. Li and R. Hartley. Five-point motion estimation made easy. In *Proc ICPR* 2006, pp. 630–633

additional references

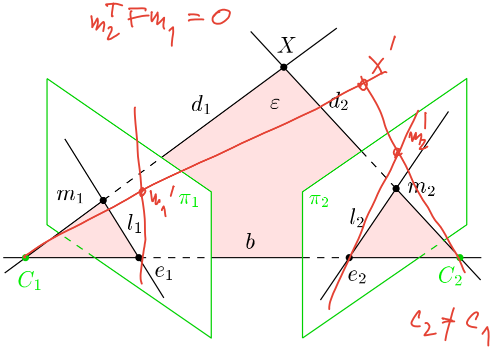
H. Longuet-Higgins. A computer algorithm for reconstructing a scene from two projections. *Nature*, 293(5828):133–135, 1981.

▶Geometric Model of a Camera Stereo Pair

$$\mathbf{P}_i = \begin{bmatrix}\mathbf{Q}_i & \mathbf{q}_i\end{bmatrix} = \mathbf{K}_i \begin{bmatrix}\mathbf{R}_i & \mathbf{t}_i\end{bmatrix} = \mathbf{K}_i\mathbf{R}_i \begin{bmatrix}\mathbf{I} & -\mathbf{C}_i\end{bmatrix} \quad i = 1, 2 \qquad \to 31$$

Epipolar geometry:

- brings constraints necessary for inter-image matching
- its parametric form encapsulates information about the relative pose of two cameras



two-camera setup

Description

- baseline b joins projection centers C_1, C_2
$$\mathbf{b} = \mathbf{C}_2 - \mathbf{C}_1$$
- epipole $e_i \in \pi_i$ is the image of C_j:
$$\underline{\mathbf{e}}_1 \simeq \mathbf{P}_1\underline{\mathbf{C}}_2, \quad \underline{\mathbf{e}}_2 \simeq \mathbf{P}_2\underline{\mathbf{C}}_1$$
- $l_i \in \pi_i$ is the image of optical ray d_j, $j \neq i$ and also the epipolar plane
$$\varepsilon = (C_2, X, C_1)$$
- l_j is the epipolar line ('epipolar') in image π_j induced by m_i in image π_i

Epipolar constraint relates $\underline{\mathbf{m}}_1$ and $\underline{\mathbf{m}}_2$: corresponding d_2, b, d_1 are coplanar

a necessary condition →88

Epipolar Geometry Example: Forward Motion

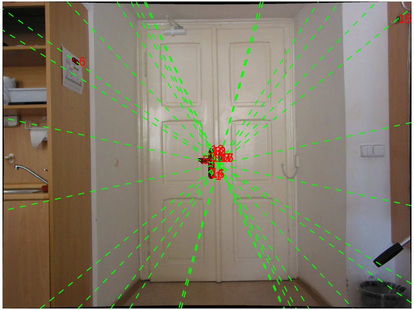


image 1

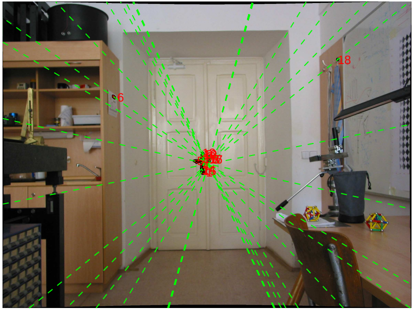
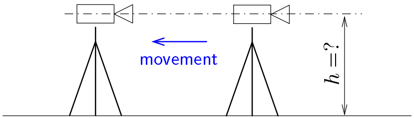


image 2

- red: correspondences
- green: epipolar line pairs per correspondence

click on the image to see their IDs
same ID in both images

Epipole is the image of the other camera's center.
How high was the camera above the floor?



▶Cross Products and Maps by Skew-Symmetric 3×3 Matrices

- There is an equivalence $\mathbf{b} \times \mathbf{m} = [\mathbf{b}]_\times \mathbf{m}$, where $[\mathbf{b}]_\times$ is a 3×3 skew-symmetric matrix

$$[\mathbf{b}]_\times = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}, \qquad \text{assuming} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

Some properties

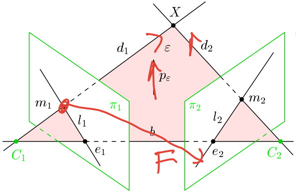
1. $[\mathbf{b}]_\times^\top = -[\mathbf{b}]_\times$ — the general antisymmetry property
2. $\mathbf{A}$ is skew-symmetric iff $\mathbf{x}^\top \mathbf{A} \mathbf{x} = 0$ for all $\mathbf{x}$ — skew-sym mtx generalizes cross products
3. $[\mathbf{b}]_\times^3 = -\|\mathbf{b}\|^2 \cdot [\mathbf{b}]_\times$
4. $\|[\mathbf{b}]_\times\|_F = \sqrt{2}\,\|\mathbf{b}\|$ — Frobenius norm ($\|\mathbf{A}\|_F = \sqrt{\mathrm{tr}(\mathbf{A}^\top\mathbf{A})} = \sqrt{\sum_{i,j} |a_{ij}|^2}$)
5. $\mathrm{rank}\,[\mathbf{b}]_\times = 2$ iff $\|\mathbf{b}\| > 0$ — check minors of $[\mathbf{b}]_\times$
6. $[\mathbf{b}]_\times \mathbf{b} = \mathbf{0}$
7. eigenvalues of $[\mathbf{b}]_\times$ are $(0, \lambda, -\lambda)$
8. for any 3×3 regular $\mathbf{B}$: $\quad \mathbf{B}^\top [\mathbf{B}\mathbf{z}]_\times \mathbf{B} = \det(\mathbf{B})\,[\mathbf{z}]_\times$ — follows from the factoring on →39
9. in particular: $\quad$ if $\mathbf{R}\mathbf{R}^\top = \mathbf{I}$ then $\boxed{[\mathbf{R}\cancel{\mathbf{b}}_{z}]_\times = \mathbf{R}[\cancel{\mathbf{b}}_{z}]_\times \mathbf{R}^\top}$

3x3: 0 ×; 1 inspection; 2; 3 ×

$b \times b = 0 \Rightarrow [b]_\times b = 0$

- ~~note that if $\mathbf{R}_b$ is rotation about $\mathbf{b}$ then $\mathbf{R}_b \mathbf{b} = \mathbf{b}$~~
- note $[\mathbf{b}]_\times$ is not a homography; it is not a rotation matrix — it is the logarithm of a rotation mtx

▶Expressing the Epipolar Constraint Algebraically: Fundamental Matrix



(d_1, d_2, b) co-planar

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

ℓ_2

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

$=$

ℓ_1

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

$m_1^T F^T m_2 = \emptyset$

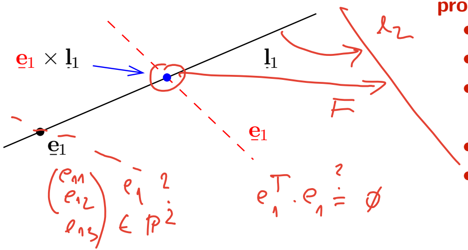
$$\mathbf{F} = (\underbrace{\mathbf{Q}_2\mathbf{Q}_1^{-1}}_{\text{epipolar homography } \mathbf{H}_e})^{-\top} [\underline{\mathbf{e}}_1]_\times = \mathbf{H}_e^{-\top} [\overbrace{\underline{\mathbf{e}}_1}^{\text{left epipole}}]_\times \overset{\to 76}{\simeq} [\overbrace{\mathbf{H}_e \underline{\mathbf{e}}_1}^{\text{right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

- 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 $\underline{\mathbf{l}}_1 \simeq \mathbf{F}^\top \underline{\mathbf{m}}_2$
- all epipolars meet at the epipole
- epipolar lines map by epipolar homography $\mathbf{H}_e^{-\top}$
- epipoles map by epipolar homography $\mathbf{H}_e$

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

- epipole $\underline{\mathbf{e}}_1$ falls in the nullspace of $\mathbf{F}$: $\mathbf{F}\underline{\mathbf{e}}_1 = \mathbf{H}_e^{-\top}[\underline{\mathbf{e}}_1]_\times\underline{\mathbf{e}}_1 = \mathbf{0}$, also $\underline{\mathbf{e}}_2^\top\mathbf{F} = \mathbf{0}$
- $\mathbf{F}$ maps points to lines and it is <u>not a homography</u>

- $\mathbf{H}_e^{-\top}$ maps epipolars to epipolars: $\underline{\mathbf{l}}_2 \simeq \mathbf{H}_e^{-\top}\underline{\mathbf{l}}_1$
- there is another useful map that does the job for epipolars: $\underline{\mathbf{l}}_2 \simeq \mathbf{F}[\underline{\mathbf{e}}_1]_\times\underline{\mathbf{l}}_1 = \mathbf{F}(\underline{\mathbf{e}}_1 \times \underline{\mathbf{l}}_1)$



proof by point/line 'transmutation' (left):

- point $\underline{\mathbf{e}}_1$ does not lie on line $\underline{\mathbf{e}}_1$ (dashed): $\underline{\mathbf{e}}_1^\top\underline{\mathbf{e}}_1 \neq 0$
- $\underline{\mathbf{e}}_1 \times \underline{\mathbf{l}}_1$ is a point on $\underline{\mathbf{l}}_1$
- $\mathbf{F}$ maps that point to $\underline{\mathbf{l}}_2$

- the composition $\mathbf{F}[\underline{\mathbf{e}}_1]_\times$ is not a homography
- usefulness: no need to decompose $\mathbf{F}$ to obtain $\mathbf{H}_e$

▶The Essential Matrix

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

$\mathbf{R}_{21}$ – 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

- the epipole is the image of the (projection center) of the other camera

$b = C_2 - C_1$

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

rank 2, homog.

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

7 DoF

$\lambda\, x_2^\top E x_1 = 0$

$$\mathbf{E} = [-\mathbf{t}_{21}]_\times\mathbf{R}_{21} = \underbrace{[\mathbf{R}_2\mathbf{b}]_\times}_{\text{baseline in Cam 2}}\mathbf{R}_{21} \overset{\to 76/9}{=} \mathbf{R}_{21}\underbrace{[\mathbf{R}_1\mathbf{b}]_\times}_{\text{baseline in Cam 1}} = \mathbf{R}_{21}[-\mathbf{R}_{21}^\top\mathbf{t}_{21}]_\times \quad \text{essential}$$

5 DoF

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

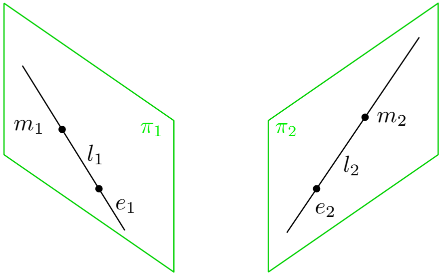
(the change of the world coordinate system by $(\mathbf{R}, \mathbf{t})$ does not change $\mathbf{E}$)

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

then

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

- the translation length $\|\mathbf{t}_{21}\|$ is <u>lost</u>, since $\mathbf{E}$ is homogeneous



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

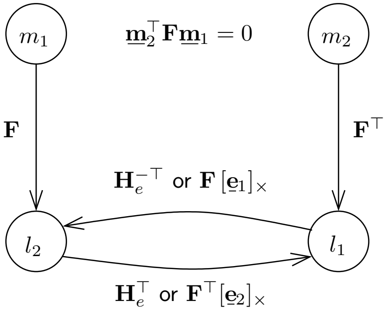
$$\underline{\mathbf{e}}_1 \simeq \mathrm{null}(\mathbf{F}), \qquad \underline{\mathbf{e}}_2 \simeq \mathrm{null}(\mathbf{F}^\top)$$

$$\underline{\mathbf{e}}_1 \simeq \mathbf{H}_e^{-1}\underline{\mathbf{e}}_2 \qquad \underline{\mathbf{e}}_2 \simeq \mathbf{H}_e\underline{\mathbf{e}}_1$$

$$\underline{\mathbf{l}}_1 \simeq \mathbf{F}^\top\underline{\mathbf{m}}_2 \qquad \underline{\mathbf{l}}_2 \simeq \mathbf{F}\underline{\mathbf{m}}_1$$

$$\underline{\mathbf{l}}_1 \simeq \mathbf{H}_e^\top\underline{\mathbf{l}}_2 \qquad \underline{\mathbf{l}}_2 \simeq \mathbf{H}_e^{-\top}\underline{\mathbf{l}}_1$$

$$\underline{\mathbf{l}}_1 \simeq \mathbf{F}^\top[\underline{\mathbf{e}}_2]_\times\underline{\mathbf{l}}_2 \qquad \underline{\mathbf{l}}_2 \simeq \mathbf{F}[\underline{\mathbf{e}}_1]_\times\underline{\mathbf{l}}_1$$



- $\mathbf{H}_e = \mathbf{Q}_2\mathbf{Q}_1^{-1}$ is the epipolar homography→79
  $\mathbf{H}_e^{-\top}$ 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

- $\mathbf{F}[\underline{\mathbf{e}}_1]_\times$ maps epipolar lines to epipolar lines but it is not a homography
- The essential matrix is the 'calibrated fundamental matrix'

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.

homogeneous

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 $\mathbf{B} = \operatorname{diag}(\lambda_1, \lambda_2, \lambda_3)$, $\mathbf{C} = \operatorname{diag}(1, 1, 0)$, $\lambda_3 > 0$
3. then $\mathbf{A} = \mathbf{U}\underbrace{\mathbf{B}\mathbf{C}}_{D}\mathbf{V}^\top = \mathbf{U}\mathbf{B}\mathbf{C}\underbrace{\mathbf{W}\mathbf{W}^\top}_{\mathbf{I}}\mathbf{V}^\top$ with $\mathbf{W}$ rotation matrix
4. we look for a rotation mtx $\mathbf{W}$ that maps $\mathbf{C}$ to a skew-symmetric $\mathbf{S}$, i.e. $\mathbf{S} = \mathbf{C}\mathbf{W}$ (if it exists)
5. then $\mathbf{W} = \begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $|\alpha| = 1$, and $\mathbf{S} = \mathbf{C}\mathbf{W} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & \alpha & 0 \\ -\alpha & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \cdots = [\mathbf{s}]_\times$, where $\mathbf{s} = (0, 0, 1)$
6. we write

$\mathbf{v}_3$ – 3rd column of $\mathbf{V}$, $\mathbf{u}_3$ – 3rd column of $\mathbf{U}$

S

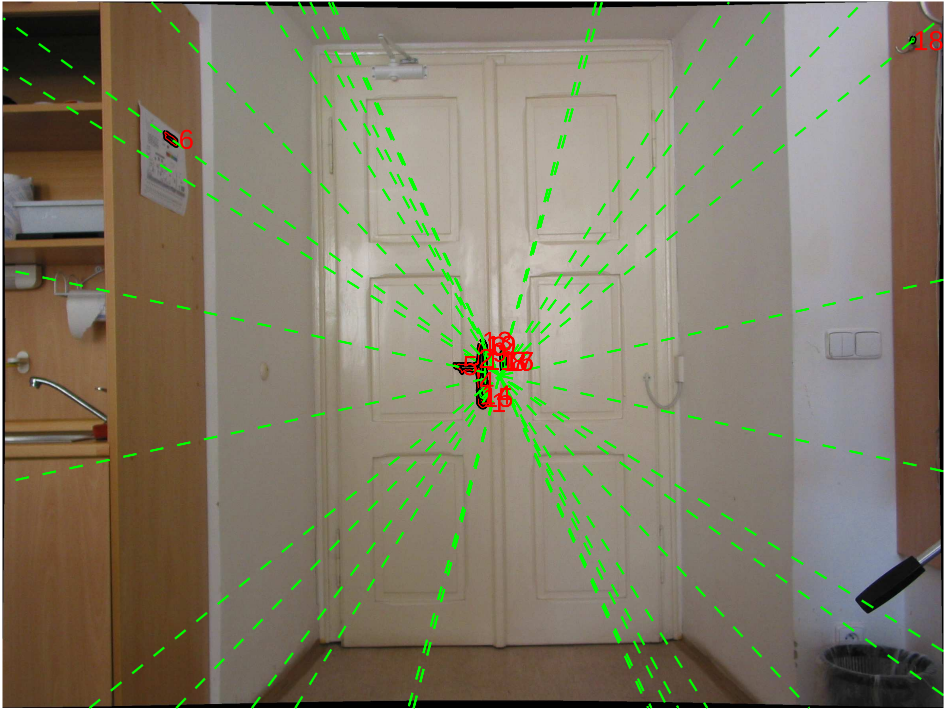
$$\mathbf{A} = \mathbf{U}\underbrace{\mathbf{B}[\mathbf{s}]_\times}_{\cancel{\mathbf{CW}}}\mathbf{W}^\top\mathbf{V}^\top \overset{\circledast\ 1}{=} \cdots = \underbrace{\mathbf{U}\mathbf{B}(\mathbf{V}\mathbf{W})^\top}_{\simeq \mathbf{H}^{-\top}} \underbrace{[\mathbf{v}_3]_\times}_{\text{3rd col } \mathbf{V}} \overset{\to 76/9}{\simeq} \underbrace{[\mathbf{H}\mathbf{v}_3]_\times}_{\simeq [\mathbf{u}_3]_\times} \mathbf{H}, \tag{12}$$

$[S]_x = CW$

7. $\mathbf{H}$ regular, $\mathbf{A}\mathbf{v}_3 = \mathbf{0}$, $\mathbf{u}_3\mathbf{A} = \mathbf{0}$ 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

Thank You


6
18
13
5
16
11
15