►Populating A Little ZOO of Minimal Geometric Problems in CV

problem	given	unknown	slide
camera resection	6 world–img correspondences $\{(X_i, m_i)\}_{i=1}^{6}$	$\mathbf{P}$	62
exterior orientation	$\mathbf{K}$, 3 world–img correspondences $\{(X_i, m_i)\}_{i=1}^{3}$	$\mathbf{R}$, $\mathbf{C}$	66
relative orientation	3 world-world correspondences $\{(X_i, Y_i)\}_{i=1}^{3}$	$\mathbf{R}$, $\mathbf{t}$	69

- camera resection and exterior orientation are similar problems in a sense:
  - we do resectioning when our camera is uncalibrated
  - we do orientation when our camera is calibrated
- relative orientation involves no camera (see next)
- more problems to come

The Relative Orientation Problem

Problem: Given two point triples (X_1, X_2, X_3) and (Y_1, Y_2, Y_3) in a general position in $\mathbf{R}^3$ such that the correspondence $X_i \leftrightarrow Y_i$ is known, determine the relative orientation $(\mathbf{R}, \mathbf{t})$ that maps $\mathbf{X}_i$ to $\mathbf{Y}_i$, i.e.

$$\mathbf{Y}_i = \mathbf{R}\mathbf{X}_i + \mathbf{t}, \quad i = 1, 2, 3\,.$$

Applies to:

- 3D scanners
- partial reconstructions from different viewpoints

Obs: Let $\bar{\mathbf{X}} = \frac{1}{3}\sum_i \mathbf{X}_i$ and analogically for $\bar{\mathbf{Y}}$. Then

$$\bar{\mathbf{Y}} = \mathbf{R}\bar{\mathbf{X}} + \mathbf{t}.$$

Therefore

$$\mathbf{Z}_i \stackrel{\text{def}}{=} (\mathbf{Y}_i - \bar{\mathbf{Y}}) = \mathbf{R}(\mathbf{X}_i - \bar{\mathbf{X}}) \stackrel{\text{def}}{=} \mathbf{R}\mathbf{W}_i$$

If all dot products are equal, $\mathbf{Z}_i^\top \mathbf{Z}_j = \mathbf{W}_i^\top \mathbf{W}_j$ for $i, j = 1, 2, 3$, we have

$$\mathbf{R}^* = \begin{bmatrix}\mathbf{W}_1 & \mathbf{W}_2 & \mathbf{W}_3\end{bmatrix}^{-1} \begin{bmatrix}\mathbf{Z}_1 & \mathbf{Z}_2 & \mathbf{Z}_3\end{bmatrix}$$

Otherwise (in practice) we setup a minimization problem

$$\mathbf{R}^* = \arg\min_{\mathbf{R}} \sum_i \|\mathbf{Z}_i - \mathbf{R}\mathbf{W}_i\|^2 \quad \text{s.t.} \quad \mathbf{R}^\top\mathbf{R} = \mathbf{I}$$

$$\min_{\mathbf{R}} \sum_i \|\mathbf{Z}_i - \mathbf{R}\mathbf{W}_i\|^2 = \min_{\mathbf{R}} \sum_i \left(\|\mathbf{Z}_i\|^2 - 2\mathbf{Z}_i^\top\mathbf{R}\mathbf{W}_i + \|\mathbf{W}_i\|^2\right) = \cdots = \max_{\mathbf{R}} \sum_i \mathbf{Z}_i^\top\mathbf{R}\mathbf{W}_i$$

cont'd (What is Linear Algebra Telling Us?)

Obs 1: Let $\mathbf{A} : \mathbf{B} = \sum_{i,j} a_{ij} b_{ij}$ be the dot-product (Frobenius inner product) over real matrices. Then

$$\mathbf{A} : \mathbf{B} = \operatorname{tr}(\mathbf{A}^\top \mathbf{B})$$

Obs 2:

$$\mathbf{Z}_i^\top \mathbf{R} \mathbf{W}_i = (\mathbf{Z}_i \mathbf{W}_i^\top) : \mathbf{R}$$

Obs 3: (cyclic property for matrix trace)

$$\operatorname{tr}(\mathbf{ABC}) = \operatorname{tr}(\mathbf{BCA})$$

Let the SVD be

$$\sum_i \mathbf{Z}_i \mathbf{W}_i^\top \stackrel{\text{def}}{=} \mathbf{M} = \mathbf{UDV}^\top$$

Then

$$\mathbf{R} : \mathbf{M} = \mathbf{R} : (\mathbf{UDV}^\top) = \operatorname{tr}(\mathbf{R}^\top \mathbf{UDV}^\top) = \operatorname{tr}(\mathbf{V}^\top \mathbf{R}^\top \mathbf{UD}) = (\mathbf{U}^\top \mathbf{RV}) : \mathbf{D}$$

cont'd: The Algorithm

We are solving

$$\mathbf{R}^* = \arg\max_{\mathbf{R}} \sum_i \mathbf{Z}_i^\top \mathbf{R} \mathbf{W}_i = \arg\max_{\mathbf{R}} \left(\mathbf{U}^\top \mathbf{R} \mathbf{V}\right) : \mathbf{D}$$

It follows that $\mathbf{U}^\top \mathbf{R} \mathbf{V}$ must be (1) diagonal, (2) orthogonal, (3) positive definite matrix. Since $\mathbf{U}$, $\mathbf{V}$ are orthogonal matrices then the solution to the problem is $\mathbf{R}^* = \mathbf{U}\mathbf{S}\mathbf{V}^\top$, where $\mathbf{S}$ is diagonal and orthogonal, i.e. one of

$$\pm\,\mathrm{diag}(1,1,1), \quad \pm\,\mathrm{diag}(1,-1,-1), \quad \pm\,\mathrm{diag}(-1,1,-1), \quad \pm\,\mathrm{diag}(-1,-1,1)$$

whichever gives $(\mathbf{R}^*)^\top \mathbf{R}^* = \mathbf{I}$

Alg:

1. Compute matrix $\mathbf{M} = \sum_i \mathbf{Z}_i \mathbf{W}_i^\top$.
2. Compute SVD $\mathbf{M} = \mathbf{U}\mathbf{D}\mathbf{V}^\top$.
3. Compute all $\mathbf{R}_k = \mathbf{U}\mathbf{S}_k\mathbf{V}^\top$ that give $\mathbf{R}_k^\top \mathbf{R}_k = \mathbf{I}$.
4. Compute $\mathbf{t}_k = \bar{\mathbf{Y}} - \mathbf{R}_k \bar{\mathbf{X}}$.

- The algorithm can be used for more than 3 points
- The P3P problem is very similar but not identical

Module IV

Computing with a Camera Pair

4.1 Camera Motions Inducing Epipolar Geometry

4.2 Estimating Fundamental Matrix from 7 Correspondences

4.3 Estimating Essential Matrix from 5 Correspondences

4.4 Triangulation: 3D Point Position from a Pair of Corresponding Points

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

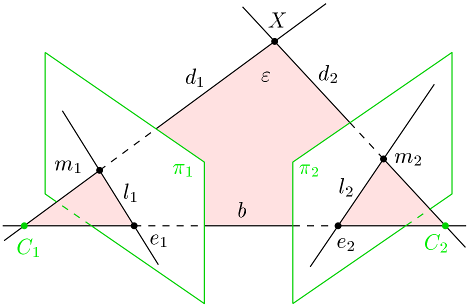
additional references

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

▶Geometric Model of a Camera Pair

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 epipolar plane

$$\varepsilon = (C_2, X, C_1)$$

- l_j is the epipolar line in image π_j induced by m_i in image π_i

Epipolar constraint: corresponding d_2, b, d_1 are coplanar   a necessary condition →86

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

→31

Epipolar Geometry Example: Forward Motion

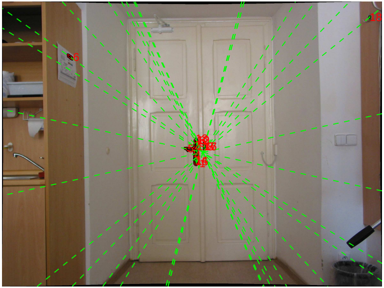


image 1

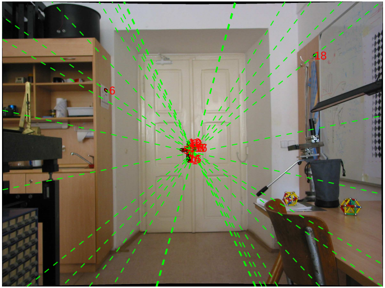
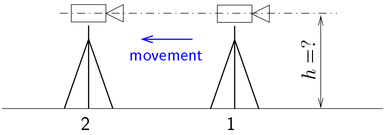


image 2

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

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

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

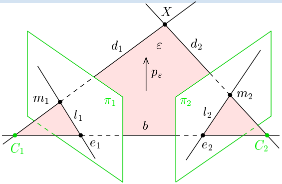
$$[\mathbf{b}]_\times = \begin{bmatrix} 0 & -b_3 & b_2 \\ b_3 & 0 & -b_1 \\ -b_2 & b_1 & 0 \end{bmatrix}, \qquad \text{assuming} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

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. $[\mathbf{b}]_\times \mathbf{b} = \mathbf{0}$
6. $\operatorname{rank} [\mathbf{b}]_\times = 2$ iff $\|\mathbf{b}\| > 0$ — check minors of $[\mathbf{b}]_\times$
7. eigenvalues of $[\mathbf{b}]_\times$ are $(0, \lambda, -\lambda)$
8. for any regular $\mathbf{B}$: $\mathbf{B}^\top [\mathbf{B}\mathbf{z}]_\times \mathbf{B} = \det \mathbf{B}\, [\mathbf{z}]_\times$ — follows from the factoring on →38
9. in particular: if $\mathbf{R}\mathbf{R}^\top = \mathbf{I}$ then $\mathbf{R}^\top [\mathbf{R}\mathbf{b}]_\times \mathbf{R} = [\mathbf{b}]_\times$

- 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 a logarithm of a rotation mtx

▸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, $\mathbf{t}_{21} = \mathbf{t}_2 - \mathbf{R}_{21}\mathbf{t}_1 = -\mathbf{R}_2\mathbf{b}$ →73

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

remember: $\mathbf{C} = -\mathbf{Q}^{-1}\mathbf{q} = -\mathbf{R}^\top\mathbf{t}$ →32 and 34

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

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

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

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

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

►The Structure and the Key Properties of the Fundamental Matrix

$$\mathbf{F} = (\underbrace{\mathbf{Q}_2\mathbf{Q}_1^{-1}}_{\text{epipolar homography } \mathbf{H}_e})^{-\top}[\mathbf{e}_1]_\times = \underbrace{\mathbf{K}_2^{-\top}\mathbf{R}_{21}\mathbf{K}_1^{\top}}_{\mathbf{H}_e^{-\top}}[\overbrace{\mathbf{e}_1}^{\text{left epipole}}]_\times \overset{\rightarrow 75}{\simeq} [\overbrace{\mathbf{H}_e\mathbf{e}_1}^{\text{right epipole}}]_\times \mathbf{H}_e = \mathbf{K}_2^{-\top}\underbrace{[-\mathbf{t}_{21}]_\times \mathbf{R}_{21}}_{\text{essential matrix } \mathbf{E}}\mathbf{K}_1^{-1}$$

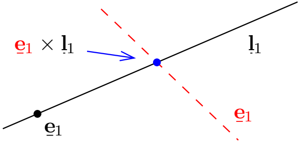
1. $\mathbf{E}$ captures relative camera pose only [Longuet-Higgins 1981]
   (the change of the world coordinate system does not change $\mathbf{E}$)

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

   then

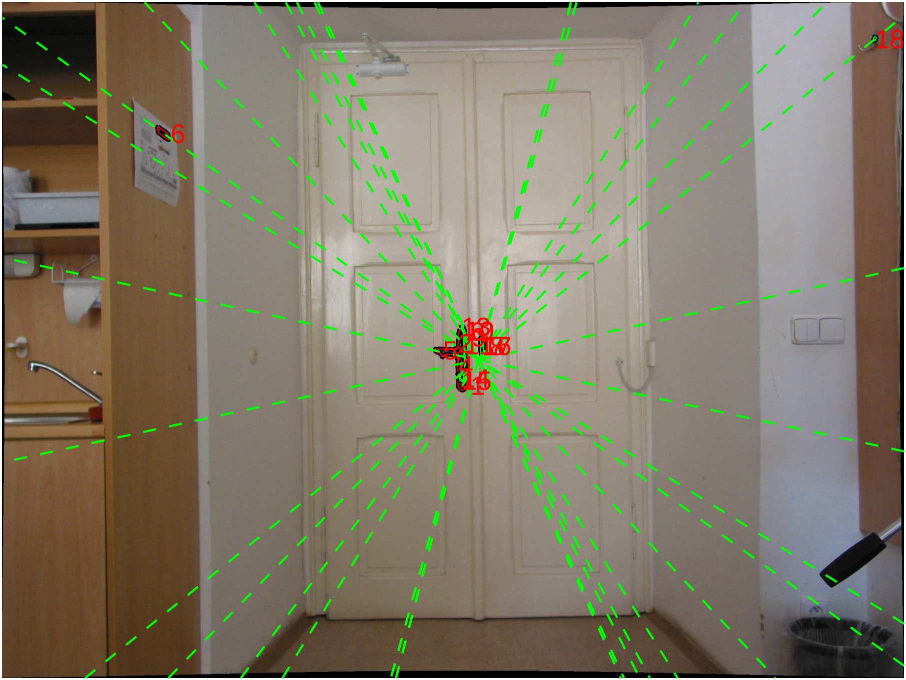
$$\mathbf{R}_{21}' = \mathbf{R}_2'\mathbf{R}_1'^{\top} = \cdots = \mathbf{R}_{21} \qquad \mathbf{t}_{21}' = \mathbf{t}_2' - \mathbf{R}_{21}'\mathbf{t}_1' = \cdots = \mathbf{t}_{21}$$

2. the translation length $\mathbf{t}_{21}$ is lost 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: $\underline{\mathbf{l}}_2 \simeq \mathbf{H}_e^{-\top}\underline{\mathbf{l}}_1$



- replacement for $\mathbf{H}_e^{-\top}$ for 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

Thank You


6
18