OPPA European Social Fund
Prague & EU: We invest in your future.

▶The Triangulation Problem

Problem: Given cameras $\mathbf{P}_1$, $\mathbf{P}_2$ and a correspondence $x \leftrightarrow y$ compute a 3D point $\mathbf{X}$ projecting to x and y

$$\lambda_1\, \underline{\mathbf{x}} = \mathbf{P}_1 \underline{\mathbf{X}}, \qquad \lambda_2\, \underline{\mathbf{y}} = \mathbf{P}_2 \underline{\mathbf{X}}, \qquad \underline{\mathbf{x}} = \begin{bmatrix} u^1 \\ v^1 \\ 1 \end{bmatrix}, \qquad \underline{\mathbf{y}} = \begin{bmatrix} u^2 \\ v^2 \\ 1 \end{bmatrix}, \qquad \mathbf{P}_i = \begin{bmatrix} (\mathbf{p}_1^i)^\top \\ (\mathbf{p}_2^i)^\top \\ (\mathbf{p}_3^i)^\top \end{bmatrix}$$

Linear triangulation method

$$u^1\, (\mathbf{p}_3^1)^\top \underline{\mathbf{X}} = (\mathbf{p}_1^1)^\top \underline{\mathbf{X}}, \qquad u^2\, (\mathbf{p}_3^2)^\top \underline{\mathbf{X}} = (\mathbf{p}_1^2)^\top \underline{\mathbf{X}},$$

$$v^1\, (\mathbf{p}_3^1)^\top \underline{\mathbf{X}} = (\mathbf{p}_2^1)^\top \underline{\mathbf{X}}, \longleftrightarrow \quad v^2\, (\mathbf{p}_3^2)^\top \underline{\mathbf{X}} = (\mathbf{p}_2^2)^\top \underline{\mathbf{X}},$$

Gives

v_1 v_2 $v_1 = v_2$

$$\mathbf{D}\underline{\mathbf{X}} = \mathbf{0}, \qquad \mathbf{D} = \begin{bmatrix} u^1\, (\mathbf{p}_3^1)^\top - (\mathbf{p}_1^1)^\top \\ v^1\, (\mathbf{p}_3^1)^\top - (\mathbf{p}_2^1)^\top \\ u^2\, (\mathbf{p}_3^2)^\top - (\mathbf{p}_1^2)^\top \\ v^2\, (\mathbf{p}_3^2)^\top - (\mathbf{p}_2^2)^\top \end{bmatrix}, \qquad \mathbf{D} \in \mathbb{R}^{4,4}, \quad \underline{\mathbf{X}} \in \mathbb{R}^4 \tag{12}$$

$$P_1 = \begin{bmatrix} p_1^{1\top} \\ p_2^{1\top} \\ p_3^{1\top} \end{bmatrix} \qquad P_2 = \begin{bmatrix} p_1^{2\top} \\ p_2^{1} \\ p_3^{1} \end{bmatrix}$$

- back-projected rays will generally not intersect due to image error, see next
- using Jack-knife (Slide 66) not recommended — sensitive to small error
- we will use SVD (Slide 86)
- but the result will not be invariant to projective frame

  replacing $\mathbf{P}_1 \mapsto \mathbf{P}_1\mathbf{H}$, $\mathbf{P}_2^{\times} \mapsto \mathbf{P}_2^{\times}\mathbf{H}$ does not always result in $\underline{\mathbf{X}} \mapsto \mathbf{H}^{-1}\underline{\mathbf{X}}$
- the homogeneous form in (12) can represent points at infinity

▶The Least-Squares Triangulation by SVD

- if $\mathbf{D}$ is full-rank we may minimize the algebraic least-squares error

$X = \arg\min_{q} \varepsilon^2(q)$

$$\varepsilon^2(\underline{\mathbf{X}}) = \|\mathbf{D}\underline{\mathbf{X}}\|^2 \quad \text{s.t.} \quad \|\underline{\mathbf{X}}\| = 1, \qquad \underline{\mathbf{X}} \in \mathbb{R}^4$$

$\lambda X \equiv X$

$[U, D, V] = \mathrm{svd}(D);$
$X = V(:, 3);$

- let $\mathbf{D}_i$ be the i-th row of $\mathbf{D}$, then

$$\|\mathbf{D}\underline{\mathbf{X}}\|^2 = \sum_{i=1}^{4} (\mathbf{D}_i\,\underline{\mathbf{X}})^2 = \sum_{i=1}^{4} \underline{\mathbf{X}}^\top \mathbf{D}_i^\top \mathbf{D}_i\,\underline{\mathbf{X}} = \underline{\mathbf{X}}^\top \mathbf{Q}\,\underline{\mathbf{X}}, \text{ where } \overset{Q}{\cancel{\underline{\mathbf{X}}}} = \sum_{i=1}^{4} \mathbf{D}_i^\top \mathbf{D}_i = \mathbf{D}^\top \mathbf{D} \in \mathbb{R}^{4,4}$$

$(D_i X)^\top (D_i X) = X^\top D_i^\top D_i X$

- we write the SVD of $\mathbf{Q}$ as $\mathbf{Q} = \sum_{j=1}^{4} \sigma_j^2\, \mathbf{u}_j \mathbf{u}_j^\top$, in which [Golub & van Loan 1996, Sec. 2.5]

$$\sigma_1^2 \geq \cdots \geq \sigma_4^2 \geq 0 \quad \text{and} \quad \mathbf{u}_l^\top \mathbf{u}_m = \begin{cases} 0 & \text{if } l \neq m \\ 1 & \text{otherwise} \end{cases}$$

- then

$$\underline{\mathbf{X}} = \arg \min_{\mathbf{q}, \|\mathbf{q}\|=1} \mathbf{q}^\top \mathbf{Q}\,\mathbf{q} = \mathbf{u}_4, \qquad \mathbf{q}^\top \mathbf{Q}\,\mathbf{q} = \sum_{j=1}^{4} \sigma_j^2\, \mathbf{q}^\top \mathbf{u}_j\, \mathbf{u}_j^\top \mathbf{q} = \sum_{j=1}^{4} \sigma_j^2\, (\mathbf{u}_j^\top \mathbf{q})^2$$

we have a sum of non-negative elements $0 \leq (\mathbf{u}_j^\top \mathbf{q})^2 \leq 1$, let $\mathbf{q} = \mathbf{u}_4 + \bar{\mathbf{q}}$ s.t. $\bar{\mathbf{q}} \perp \mathbf{u}_4$, then

$$\mathbf{q}^\top \mathbf{Q}\,\mathbf{q} = \sigma_4^2 + \sum_{j=1}^{\cancel{4}\,3} \sigma_j^2\, (\mathbf{u}_j^\top \bar{\mathbf{q}})^2 \geq \sigma_4^2$$

▶cont'd

- if $\sigma_4 \ll \sigma_3$, there is a unique solution $\mathbf{X} = \mathbf{u}_4$ with residual error $(\mathbf{D}\,\mathbf{X})^2 = \sigma_4^2$
  the quality (conditioning) of the solution may be expressed as $q = \sigma_3/\sigma_4$ (greater is better)

Matlab code for the least-squares solver:

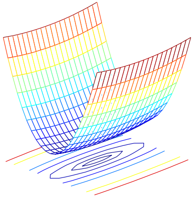
svd (2)

```
[U,O,V] = svd(D);
X = V(:,end);
q = O(3,3)/O(4,4);
```

⊛ P1; 2pt: Why did we decompose $\mathbf{D}$ and not $\mathbf{Q} = \mathbf{D}^\top\mathbf{D}$? Could we use QR decomposition instead of SVD?

▶Numerical Conditioning

- The equation $\mathbf{D}\underline{\mathbf{X}} = \mathbf{0}$ in (12) may be ill-conditioned for numerical computation, which results in a poor estimate for $\underline{\mathbf{X}}$.



Why: on a row of $\mathbf{D}$ there are big entries together with small entries, e.g. of orders   projection centers in mm, image points in px

$$\begin{bmatrix} 10^3 & 0 & 10^3 & 10^6 \\ 0 & 10^3 & 10^3 & 10^6 \\ 10^3 & 0 & 10^3 & 10^6 \\ 0 & 10^3 & 10^3 & 10^6 \end{bmatrix}$$

Quick fix:

1. re-scale the problem by a regular diagonal conditioning matrix $\mathbf{S} \in \mathbb{R}^{4,4}$

$$\mathbf{0} = \mathbf{D}\,\mathbf{q} = \underbrace{\mathbf{D}\,\mathbf{S}}_{\bar{D}}\,\underbrace{\mathbf{S}^{-1}\mathbf{q}}_{\bar{q}} = \bar{\mathbf{D}}\,\bar{\mathbf{q}} \qquad \bar{q} = S^{-1} q \rightarrow q = S\bar{q}$$

   choose $\mathbf{S}$ to make the entries in $\hat{\mathbf{D}}$ all smaller than unity in absolute value:

$$\mathbf{S} = \operatorname{diag}(10^{-3}, 10^{-3}, 10^{-3}, 10^{-6})$$

```
S = diag(1./max(max(abs(D)),1))
```

2. solve for $\bar{\mathbf{q}}$ as before
3. get the final solution as $\mathbf{q} = \mathbf{S}\,\bar{\mathbf{q}}$

- when SVD is used in camera resectioning, conditioning is essential for success → Slide 65

▶Back to Triangulation: The Golden Standard Method

We are given $\mathbf{P}_1$, $\mathbf{P}_2$ and a single correspondence $x \leftrightarrow y$ and we look for 3D point $\mathbf{X}$ projecting to x and y. → Slide 85

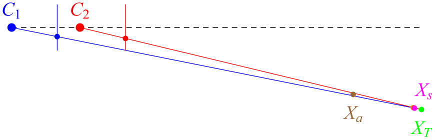
Idea:

1. compute $\mathbf{F}$ from $\mathbf{P}_1$, $\mathbf{P}_2$, e.g. $\mathbf{F} = (\mathbf{Q}_1\mathbf{Q}_2^{-1})^\top [\mathbf{q}_1 - (\mathbf{Q}_1\mathbf{Q}_2^{-1})\mathbf{q}_2]_\times$
2. correct measurement by linear estimate of the correction vector

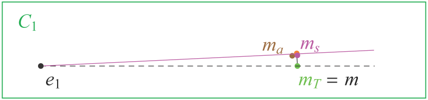
$$\begin{bmatrix} \hat{u}^1 \\ \hat{v}^1 \\ \hat{u}^2 \\ \hat{v}^2 \end{bmatrix} \approx \begin{bmatrix} u^1 \\ v^1 \\ u^2 \\ v^2 \end{bmatrix} - \frac{\varepsilon}{\|\mathbf{J}\|^2}\mathbf{J}^\top = \begin{bmatrix} u^1 \\ v^1 \\ u^2 \\ v^2 \end{bmatrix} - \frac{\underline{\mathbf{y}}^\top \mathbf{F}\underline{\mathbf{x}}}{\|\mathbf{S}\mathbf{F}\underline{\mathbf{x}}\|^2 + \|\mathbf{S}\mathbf{F}^\top\underline{\mathbf{y}}\|^2} \begin{bmatrix} (\mathbf{F}_1)^\top \mathbf{y} \\ (\mathbf{F}_2)^\top \mathbf{y} \\ (\mathbf{F}^1)^\top \mathbf{x} \\ (\mathbf{F}^2)^\top \mathbf{x} \end{bmatrix}$$

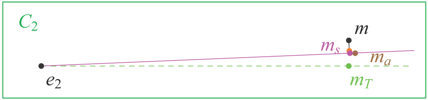
3. use the SVD algorithm with numerical conditioning → Slide 86

Ex (cont'd from Slide 89):



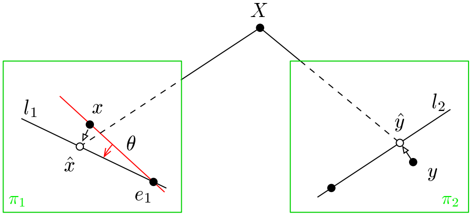
X_T – noiseless ground truth position
● – reprojection error minimizer
X_s – Sampson-corrected algebraic error minimizer
X_a – algebraic error minimizer
m – measurement (m_T with noise in v^2)





Optimal Triangulation for the Geeks

- detected image points x, y do not satisfy epipolar geometry exactly
- as a result optical rays do not intersect in space, we must correct the image points to $\hat{x}$, $\hat{y}$ first



1. given epipolar line l_1 and l_2, $\underline{\mathbf{l}}_2 \simeq \mathbf{F}[\underline{\mathbf{e}}_1]_\times \underline{\mathbf{l}}_1$ the $\hat{x}$, $\hat{y}$ are the closest points on l_1, l_2
2. parameterize all possible l_1 by θ
   - find θ after translating $\underline{\mathbf{x}}$, $\underline{\mathbf{y}}$ to $(0,0,1)$, rotating the epipoles to $(1,0,f_1)$, $(1,0,f_2)$, and parameterising $\underline{\mathbf{l}}_1 = (0,\theta,1) \times (1,0,f_1)$
3. minimise the error
$$\theta^* = \arg\min_\theta d^2\big(x, l_1(\theta)\big) + d^2\big(y, l_2(\theta)\big)$$
the problem reduces to 6-th degree polynomial root finding, see [H&Z, Sec 12.5.2]
4. compute $\hat{x}$, $\hat{y}$ and triangulate using the linear method on Slide 85

- the midpoint of the common perpendicular to both optical rays gives about 50% greater error in 3D
- a fully optimal procedure requires error re-definition in order to get the most probable $\hat{x}$, $\hat{y}$

►We Have Added to The ZOO

Continuation from Slide 71

problem	given	unknown	slide
resectioning	6 world–img correspondences $\{(X_i, m_i)\}_{i=1}^{6}$	$\mathbf{P}$	65
exterior orientation	$\mathbf{K}$, 3 world–img correspondences $\{(X_i, m_i)\}_{i=1}^{3}$	$\mathbf{R}$, $\mathbf{C}$	69
fundamental matrix	7 img–img correspondences $\{(m_i, m_i')\}_{i=1}^{7}$	$\mathbf{F}$	81
relative orientation	$\mathbf{K}$, 5 img–img correspondences $\{(m_i, m_i')\}_{i=1}^{5}$	$\mathbf{R}$, $\mathbf{t}$	84
triangulation	1 img–img correspondence (m_i, m_i')	X	85

A bigger ZOO at `http://cmp.felk.cvut.cz/minimal/`

calibrated problems

- have fewer degenerate configurations
- can do with fewer points (good for geometry proposal generators → Slide 113)

- algebraic error optimization (with SVD) makes sense in resectioning and triangulation only
- but it is not the best method; we will now focus on 'optimizing optimally'

Part V

Optimization for 3D Vision

covered by

[1] [H&Z] Secs: 11.4, 11.6, 4.7

[2] Fischler, M.A. and Bolles, R.C . Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24(6):381–395, 1981

additional references

P. D. Sampson. Fitting conic sections to 'very scattered' data: An iterative refinement of the Bookstein algorithm. *Computer Vision, Graphics, and Image Processing*, 18:97–108, 1982.

O. Chum, J. Matas, and J. Kittler. Locally optimized RANSAC. In *Proc DAGM*, LNCS 2781:236–243. Springer-Verlag, 2003.

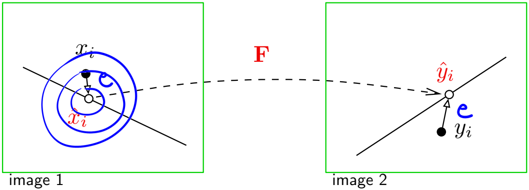
O. Chum, T. Werner, and J. Matas. Epipolar geometry estimation via RANSAC benefits from the oriented epipolar constraint. In *Proc ICPR*, vol 1:112–115, 2004.

►The Concept of Error for Epipolar Geometry

Problem: Given at least 8 corresponding points $x_i \leftrightarrow y_j$ in a general position, estimate the most likely (or most probable) fundamental matrix $\mathbf{F}$.

measurements $\quad \mathbf{x}_i = (u_i^1,\, v_i^1), \quad \mathbf{y}_i = (u_i^2,\, v_i^2), \qquad i = 1, 2, \dots, k, \quad k \geq 8$

$p(x_i) \sim N(\hat{x}, \sigma^2)$



- detected points x_i, y_i; the correspondence set is $S = \{(x_i,\, y_i)\}_{i=1}^{k}$
- corrected points $\hat{x}_i$, $\hat{y}_i$; the set is $\hat{S} = \{(\hat{x}_i,\, \hat{y}_i)\}_{i=1}^{k}$
- corrected points satisfy the epipolar geometry exactly $\underline{\hat{\mathbf{y}}}_i^\top \mathbf{F}\, \underline{\hat{\mathbf{x}}}_i = 0$ for all $i = 1, \dots, k$
- small correction is more probable
- ok, but we need to choose a definite error function for optimization that is tractable

- the solution for calibrated cameras (unknown $\mathbf{E}$) is essentially the same and is not mentioned here explicitly

►cont'd

- Let $V(\cdot)$ be a positive semi-definite 'energy function'
- e.g., per correspondence,

$- \log p$

$$V_i(x_i, y_i \mid \hat{x}_i, \hat{y}_i, \mathbf{F}) = \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 + \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2 \tag{13}$$

- the total (negative) log-likelihood (of all data) then is

$$L(S \mid \hat{S}, \mathbf{F}) = \sum_{i=1}^{k} V_i(x_i, y_i \mid \hat{x}_i, \hat{y}_i, \mathbf{F})$$

- and the optimization problem is

$$(\hat{S}^*, \mathbf{F}^*) = \arg \min_{\substack{\mathbf{F} \\ \operatorname{rank} \mathbf{F} = 2}} \; \min_{\substack{\hat{S} \\ \underline{\hat{\mathbf{y}}_i^\top \mathbf{F} \hat{\mathbf{x}}_i} = 0}} \sum_{i=1}^{k} V_i(x_i, y_i \mid \hat{x}_i, \hat{y}_i, \mathbf{F}) \tag{14}$$

$4k+7$

we mention 3 approaches

1. direct optimization of 'geometric error' over all variables $\hat{S}$, $\mathbf{F}$ Slide 95
2. approximate minimization of $L(S \mid \hat{S}, \mathbf{F})$ over $\hat{S}$ followed by minimization over $\mathbf{F}$ Slide 96
3. marginalization of $L(S, \hat{S} \mid \mathbf{F})$ over $\hat{S}$ followed by minimization over $\mathbf{F}$

Method 1: Geometric Error Optimization

- we need to encode the constraints $\hat{\underline{\mathbf{y}}}_i\, \mathbf{F}\, \hat{\underline{\mathbf{x}}}_i = 0$, $\operatorname{rank} \mathbf{F} = 2$
- idea: reconstruct 3D point via equivalent projection matrices and use reprojection error
- equivalent projection matrices are see [H&Z,Sec. 9.5] for complete characterization

$(\hat{x}, \hat{y}) \to X$

$$\mathbf{P}_1 = \begin{bmatrix} \mathbf{I} & \mathbf{0} \end{bmatrix}, \quad \mathbf{P}_2 = \begin{bmatrix} [\underline{\mathbf{e}}_2]_\times \mathbf{F} + \underline{\mathbf{e}}_2 \underline{\mathbf{e}}_1^\top & \underline{\mathbf{e}}_2 \end{bmatrix}$$

⊛ H3; 2pt: Verify that $\mathbf{F}$ is a f.m. of $\mathbf{P}_1$, $\mathbf{P}_2$, for instance that $\mathbf{F} \simeq \mathbf{Q}_2^{-\top} \mathbf{Q}_1^\top [\mathbf{e}_1]_\times$

1. compute $\mathbf{F}^{(0)}$ by the 7-point algorithm — Slide 81
2. construct camera $\mathbf{P}_2^{(0)}$ from $\mathbf{F}^{(0)}$
3. triangulate 3D points $\hat{X}_i^{(0)}$ from correspondences (x_i, y_i) for all $i = 1, \ldots, k$ — Slide 85
4. express the energy function as reprojection error

$$W_i(x_i, y_i \mid \hat{X}_i, \mathbf{P}_2) = \|\mathbf{x}_i - \hat{\mathbf{x}}_i\|^2 + \|\mathbf{y}_i - \hat{\mathbf{y}}_i\|^2 \quad \text{where} \quad \hat{\underline{\mathbf{x}}}_i \simeq \mathbf{P}_1 \hat{\underline{\mathbf{X}}}_i, \ \hat{\underline{\mathbf{y}}}_i \simeq \mathbf{P}_2(\mathbf{F})\, \hat{\underline{\mathbf{X}}}_i$$

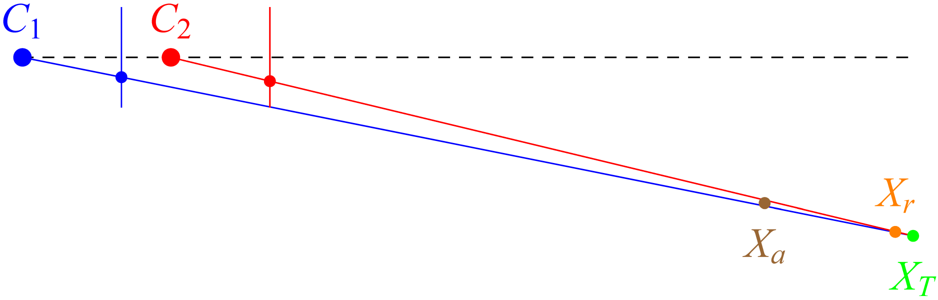
5. starting from $\mathbf{P}_2^{(0)}$, $\hat{X}^{(0)}$ minimize

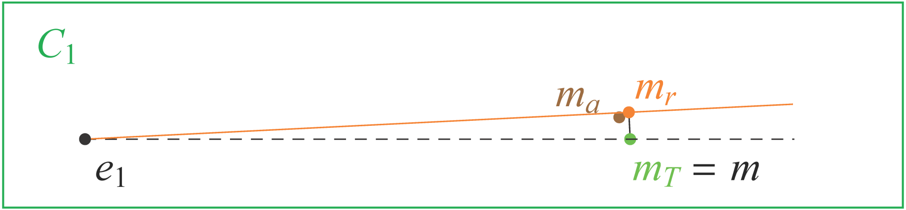
$$(\hat{X}^*, \mathbf{P}_2^*) = \arg\min_{\mathbf{P}_2,\, \hat{X}} \sum_{i=1}^{k} W_i(x_i, y_i \mid \hat{X}_i, \mathbf{P}_2)$$

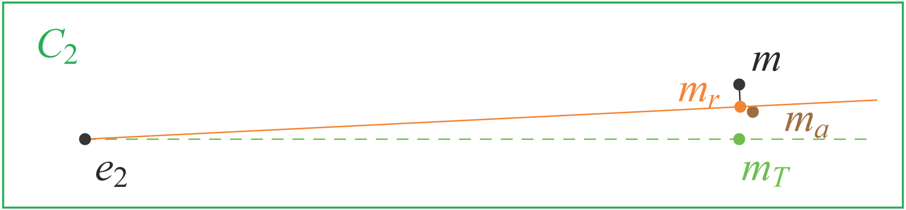
6. compute $\mathbf{F}$ from $\mathbf{P}_1$, $\mathbf{P}_2^*$

- $3k + 12$ 'parameters' to be found: latent: $\hat{\mathbf{X}}_i$, for all i (correspondences!), non-latent: $\mathbf{P}_2$
- there are pitfalls; this is essentially bundle adjustment; we will return to this later — Slide 139

Thank You


C_1
C_2
X_r
X_a
X_T


C_1
m_a
m_r
e_1
$m_T = m$


C_2
m
m_r
m_a
e_2
m_T

OPPA European Social Fund
Prague & EU: We invest in your future.