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 15, 2022

Open Informatics Master's Course

▶The Nine Elements of a Data-Driven MH Sampler

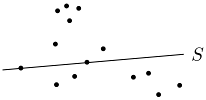
data-driven = proposals $q(S \mid C_t)$ are derived from data

Then

1. **primitives** = elementary measurements
   - points in line fitting
   - matches in epipolar geometry or homography estimation

2. **configuration** = s-tuple of primitives

minimal subsets necessary for parameter estimate



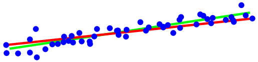
the minimization will then be over a discrete set:
   - of point pairs in line fitting (left)
   - of match 7-tuples in epipolar geometry estimation

3. a map from configuration C to parameters $\boldsymbol{\theta} = \boldsymbol{\theta}(C)$ by solving the **minimal problem**
   - line parameters $\mathbf{n}$ from two points $(\mathbf{x}^1, \mathbf{x}^2) \mapsto \mathbf{n}$
   - fundamental matrix $\mathbf{F}$ from seven matches $\{(\mathbf{x}_i^1, \mathbf{x}_i^2)\}_{i=1:7} \mapsto \mathbf{F}$
   - homography $\mathbf{H}$ from four matches, etc $\{(\mathbf{x}_i^1, \mathbf{x}_i^2)\}_{i=1:4} \mapsto \mathbf{H}$

4. **target likelihood** $p(E, D \mid \boldsymbol{\theta}(C))$ is represented by $\pi(C)$
   - can use log-likelihood: then it is the sum of robust errors $\hat{V}(e_{ij})$ given $\mathbf{F}$ (27)
     - robustified point distance from the line $\boldsymbol{\theta} = \mathbf{n}$
     - robustified Sampson error for $\boldsymbol{\theta} = \mathbf{F}$, etc
   - posterior likelihood $p(E, D \mid \boldsymbol{\theta})p(\boldsymbol{\theta})$ can be used

MAPSAC ($\pi(S)$ includes the prior)

5. parameter distribution follows the **empirical distribution** of the s-tuples of primitives. Since the proposal is done via the minimal problem solver, it is 'data-driven',



- pairs of points define line distribution $p(\mathbf{n} \mid X)$ (left)
- random correspondence 7-tuples define epipolar geometry distribution $p(\mathbf{F} \mid M)$

6. **proposal distribution** $q(\cdot)$ is just a constant(!) distribution of the s-tuples:
   a) q uniform, independent $q(S \mid C_t) = q(S) = \binom{mn}{s}^{-1}$, then $a = \min\left\{1, \frac{p(S)}{p(C_t)}\right\}$
   b) q dependent on descriptor similarity — PROSAC (similar pairs are proposed more often)
   c) q dependent on the current configuration C_t — not constant — e.g. 'not far from C_t'

7. (optional) hard **inlier/outlier discrimination** by the threshold (28)

$$\hat{V}(e_{ij}) < e_T, \qquad e_T = \sigma_1 \sqrt{-\log t^2}$$

8. **local optimization** from promising proposals

- can use the hard inliers or just the robust error (27) — more expensive but more stable
- cannot be used to replace C_t — it would violate 'detailed balance' required for the MH scheme

9. **stopping** based on the probability of proposing an all-inlier configuration →126

▶Data-Driven Sampler Stopping

- The number of proposals N needed to hit the "true parameters" = an all-inlier configuration:

  this will tell us nothing about the accuracy of the result

P ...probability that the last proposal is an all-inlier — $1-P$ ...all previous N proposals contained outliers

ε ...the fraction of inliers among primitives, $\varepsilon \leq 1$

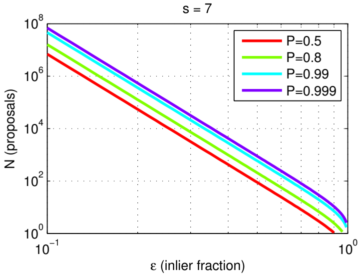
s ...No. of primitives in a minimal configuration — 2 in line fitting, 7 in 7-point algorithm, 4 in homography fitting,...

$$N \geq \frac{\log(1-P)}{\log(1-\varepsilon^s)}$$

- ε^s ...proposal is all-inlier
- $1-\varepsilon^s$ ...proposal contains at least one outlier
- $(1-\varepsilon^s)^N$ ...N previous proposals contained an outlier $= 1-P$

N for $s=7$

	P	
ε	0.8	0.99
0.5	205	590
0.2	$1.3{\cdot}10^5$	$3.5{\cdot}10^5$
0.1	$1.6{\cdot}10^7$	$4.6{\cdot}10^7$



- N can be re-estimated using the current estimate for ε (if there is LO, then after LO)

  the quasi-posterior estimate for ε is the average over all samples generated so far
- this shows we have a good reason to limit all possible matches to tentative matches only
- for $\varepsilon \to 0$ we gain nothing over the standard MH-sampler stopping rule — not covered in this course

▶Stripping MH Down To Get RANSAC [Fischler & Bolles 1981]

- when we are interested in the best config only... and we need fast data exploration...
- ... then Steps 2–4 below make no difference when waiting for the best sample configuration:

From sampling to RANSACing

1. ~~given C_t,~~ draw a random sample S from ~~$q(S \mid C_t)$~~ $q(S)$

   independent sampling
   no use of information from C_t

2. ~~compute acceptance probability~~

$$a = \min\left\{1,\ \frac{\pi(S)}{\pi(C_t)} \cdot \frac{q(C_t \mid S)}{q(S \mid C_t)}\right\}$$

3. ~~draw a random number u from unit-interval uniform distribution $\mathrm{U}_{0,1}$~~
4. ~~if $u \le a$ then $C_{t+1} := S$ else $C_{t+1} := C_t$~~
5. if $\pi(S) > \pi(C_{\text{best}})$ then remember $C_{\text{best}} := S$

- this is the ‘stupid’ Method 2 from →119
- it has a good overall exploration but slow convergence in the vicinity of a mode where C_t could serve as an attractor
- getting a good accuracy configuration might take very long this way
- (possibly robust) ‘local optimization’ necessary for reasonable performance
- unlike the full sampler, it cannot use the past generated configurations to estimate any parameters

The Opposite End: The Power of MH Sampler

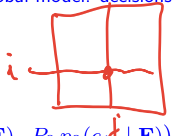
By marginalization in (23) we have lost constraints on M (e.g. uniqueness). One can choose a better model when not marginalizing:

$$\pi(M, \mathbf{F}, E, D) = \underbrace{p(E \mid M, \mathbf{F})}_{\text{reprojection error}} \cdot \underbrace{p(D \mid M)}_{\text{similarity}} \cdot \underbrace{p(\mathbf{F})}_{\text{prior}} \cdot \underbrace{P(M)}_{\text{constraints}}$$

this is a global model: decisions on m_{ij} are no longer independent!

In the MH scheme

- one can work with full $p(M, \mathbf{F} \mid E, D)$, then configuration $C = M$ — $\mathbf{F}$ computable from M



  - explicit labeling m_{ij} can be done by, e.g. sampling from

$$q(m_{ij} \mid \mathbf{F}) \sim \big((1 - P_0)\, p_1(e_{ij} \mid \mathbf{F}),\ P_0\, p_0(e_{ij} \mid \mathbf{F})\big)$$

    when $P(M)$ uniform then always accepted, $a = 1$ — ⊛ derive
  - we can compute the posterior probability of each match $p(m_{ij})$ by histogramming m_{ij} from the sequence $\{C_i\}$
  - local optimization can then use explicit inliers and $p(m_{ij})$
- error can be estimated for the elements of $\mathbf{F}$ from the sequence $\{C_i\}$ — does not work in RANSAC
- large error indicates problem degeneracy — this is not directly available in RANSAC
- good conditioning is not a requirement — we work with the entire distribution $p(\mathbf{F})$
- one can find the most probable number of models (epipolar geometries, homographies, ...) — by reversible jump MCMC
  if there are multiple models explaning data, RANSAC will return one of them randomly

Example: MH Sampling for a More Complex Problem

Task: Find two vanishing points from line segments detected in input image. Principal point is known, square pixel.

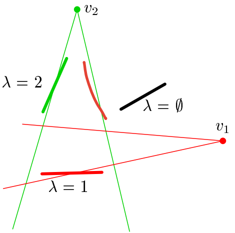
click for video

simplifications

- vanishing points restricted to the set of all pairwise segment intersections
- mother lines fixed by segment centroid, then θ_L uniquely given by λ_i, and the configuration is

$$C = \{v_1, v_2, \Lambda\}$$

- primitives = line segments
- latent variables
  1. each line has a vanishing point label $\lambda_i \in \{\emptyset, 1, 2\}$, $\emptyset$ = outlier
  2. 'mother line' parameters θ_L (they pass through their vanishing points)
- explicit variables
  1. two unknown vanishing points v_1, v_2
- marginal proposals (v_i fixed, v_j proposed)
- minimal configuration $s = 2$



$$\arg \min_{v_1, v_2, \Lambda, \theta_L} V(v_1, v_2, \Lambda, \theta_L)$$

- blue lines point away from the vanishing points
- proposal acceptance: 20%
- ca. 150 iterations to a good solution

Module VI

3D Structure and Camera Motion

6.1 Reconstructing Camera System: From Triples and from Pairs

6.2 Bundle Adjustment

covered by

[1] [H&Z] Secs: 9.5.3, 10.1, 10.2, 10.3, 12.1, 12.2, 12.4, 12.5, 18.1

[2] Triggs, B. et al. Bundle Adjustment—A Modern Synthesis. In *Proc ICCV Workshop on Vision Algorithms.* Springer-Verlag. pp. 298–372, 1999.

additional references

D. Martinec and T. Pajdla. Robust Rotation and Translation Estimation in Multiview Reconstruction. In *Proc CVPR*, 2007

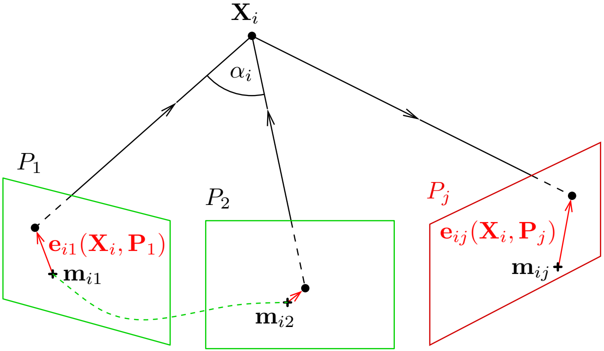
M. I. A. Lourakis and A. A. Argyros. SBA: A Software Package for Generic Sparse Bundle Adjustment. ACM Trans Math Software 36(1):1–30, 2009.

▶Reconstructing Camera System by Gluing Camera Triples

Given: Calibration matrices $\mathbf{K}_j$ and tentative correspondences per camera triples.

Initialization

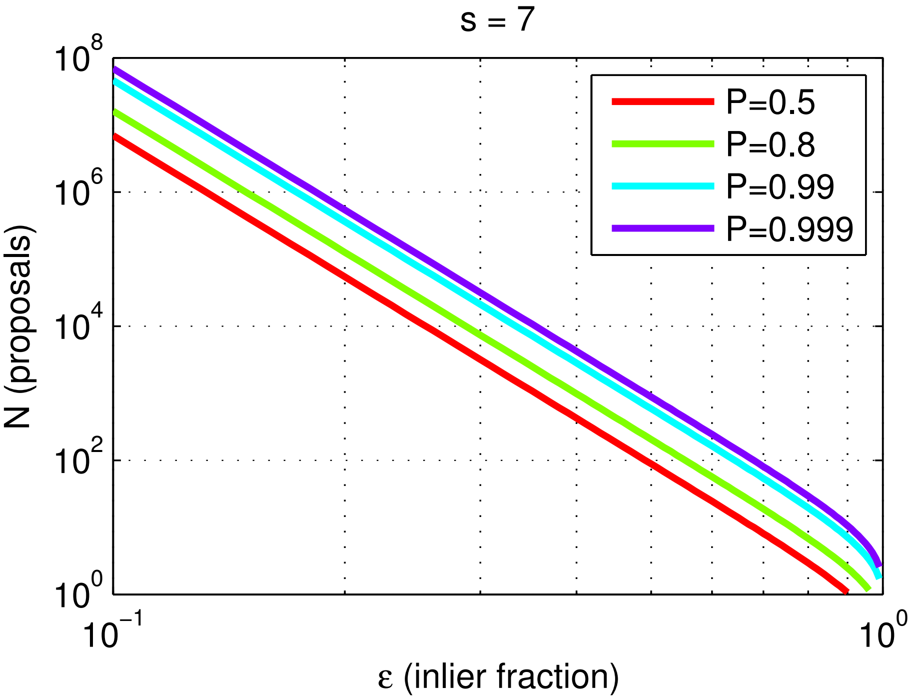
1. initialize camera cluster $\mathcal{C}$ with a pair P_1, P_2
2. find essential matrix $\mathbf{E}_{12}$ and matches M_{12} by the 5-point algorithm →88
3. construct camera pair
$$\mathbf{P}_1 = \mathbf{K}_1 \begin{bmatrix}\mathbf{I} & \mathbf{0}\end{bmatrix}, \ \mathbf{P}_2 = \mathbf{K}_2 \begin{bmatrix}\mathbf{R} & \mathbf{t}\end{bmatrix}$$
4. triangulate $\{X_i\}$ per match from M_{12} →106
5. initialize point cloud $\mathcal{X}$ with $\{X_i\}$ satisfying chirality constraint $z_i > 0$ and apical angle constraint $|\alpha_i| > \alpha_T$



Attaching camera $P_j \notin \mathcal{C}$

1. select points $\mathcal{X}_j$ from $\mathcal{X}$ that have matches to P_j
2. estimate $\mathbf{P}_j$ using $\mathcal{X}_j$, RANSAC with the 3-pt alg. (P3P), projection errors $\mathbf{e}_{ij}$ in $\mathcal{X}_j$ →66
3. reconstruct 3D points from all tentative matches from P_j to all P_l, $l \neq k$ that are not in $\mathcal{X}$
4. filter them by the chirality and apical angle constraints and add them to $\mathcal{X}$
5. add P_j to $\mathcal{C}$
6. perform bundle adjustment on $\mathcal{X}$ and $\mathcal{C}$ coming next →139

Thank You


s = 7
N (proposals)
ε (inlier fraction)
P=0.5
P=0.8
P=0.99
P=0.999
10^8
10^6
10^4
10^2
10^0
10^{-1}
10^0