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 24, 2020

Open Informatics Master's Course

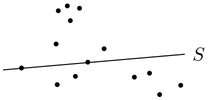
►The Nine Elements of a Data-Driven MH Sampler

data-driven = proposals 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 be over a discrete set:

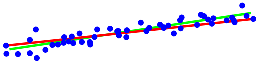
- of point pairs in line fitting (left)
- of match 7-tuples in epipolar geometry estimation

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
   - fundamental matrix $\mathbf{F}$ from seven matches
   - homography $\mathbf{H}$ from four matches, etc

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 $\mathbf{F}$ (26)
     - 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)

▶cont'd

5. parameter distribution follows the **empirical distribution** of s-tuples. 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 $q(\mathbf{F} \mid M)$

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 e.g. 'not far from C_t'

7. (optional) hard **inlier/outlier discrimination** by the threshold (27)

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

8. **local optimization** from promising proposals
   - can use the hard inliers or just the robust error (26) (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 →123

▶Data-Driven Sampler Stopping

- The number of proposals N needed to hit the "true parameters" $=$ an all-inlier config?

  this will tell us nothing about the accuracy of the result

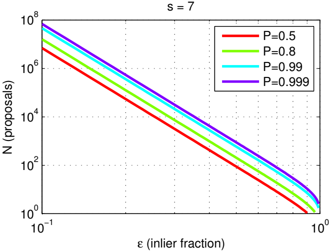
P ...probability that at least one proposal is all-inlier — $1-P$ ...all previous N proposals were bad
ε ...the fraction of inliers among primitives, $\varepsilon \leq 1$
s ...minimal configuration size — 2 in line fitting, 7 in 7-point algorithm, 4 in homography fitting,...

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

- ε^s ...proposal does not contain an outlier
- $1-\varepsilon^s$ ...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

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

- when we are interested in the best config only…and we need fast data exploration…

Simplified sampling procedure

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~~

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

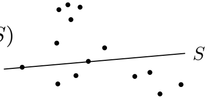
3. ~~draw a random number u from unit-interval uniform distribution $\mathrm{U}_{0,1}$~~
4. ~~if $u \leq a$ then $C_{t+1} := S$ else $C_{t+1} := C_t$~~
5. if $\pi(S) > \pi(C_{\text{best}})$ then remember $C_{\text{best}} := S$

Steps 2–4 make no difference when waiting for the best sample configuration

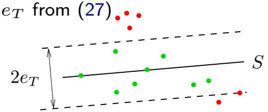
- …but getting a good accuracy configuration might take very long this way
- good overall exploration but slow convergence in the vicinity of a mode where C_t could serve as an attractor
- cannot use the past generated configurations to estimate any parameters
- we will fix these problems by (possibly robust) ‘local optimization’

►RANSAC with Local Optimization and Early Stopping

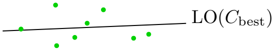
1. initialize the best configuration as empty $C_{\text{best}} := \emptyset$ and time $t := 0$
2. estimate the number of needed proposals as $N := \binom{n}{s}$ n – No. of primitives, s – minimal config size
3. while $t \le N$:
   a) propose a minimal random config S of size s from $q(S)$



   b) if $\pi(S) > \pi(C_{\text{best}})$ then
      i) update the best config $C_{\text{best}} := S$ $\pi(S)$ marginalized as in (26); $\pi(S)$ includes a prior $\Rightarrow$ MAP
      ii) threshold-out inliers using e_T from (27)



      iii) start local optimization from the inliers of C_{best} LM optimization with robustified (→114) Sampson error possibly weighted by posterior $\pi(m_{ij})$ [Chum et al. 2003]



      iv) update C_{best}, update inliers using (27), re-estimate N from inlier counts →123 for derivation

$$N = \frac{\log(1-P)}{\log(1-\varepsilon^s)}, \quad \varepsilon = \frac{|\,\text{inliers}(C_{\text{best}})|}{m\,n},$$

   c) $t := t + 1$
4. output C_{best}

- see MPV course for RANSAC details see also [Fischler & Bolles 1981], [25 years of RANSAC]

Example Matching Results for the 7-point Algorithm with RANSAC

input images | interest points (ca. 3600) | tentative corresp. (416) | matching (340)

- notice some wrong matches (they have wrong depth, even negative)
- they cannot be rejected without additional constraints or scene knowledge
- without local optimization the minimization is over a <u>discrete set</u> of epipolar geometries proposable from 7-tuples

Beyond RANSAC

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

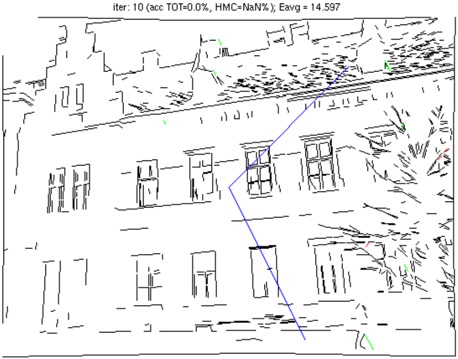
    $$q(m_{ij} \mid \mathbf{F}) \sim \big((1 - P_0)\, p_1(e_{ij} \mid \mathbf{F}),\; P_0\, p_0(e_{ij} \mid \mathbf{F})\big)$$

    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 $\{C_i\}$
  - local optimization can then use explicit inliers and $p(m_{ij})$
- error can be estimated for elements of $\mathbf{F}$ from $\{C_i\}$ — 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 epipolar geometries (homographies or other models) — by reversible jump MCMC and Bayesian model selection

if there are multiple models explaining 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.



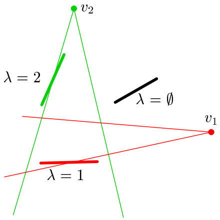
video

- primitives = line segments
- latent variables
  1. each line has a vanishing point label $\lambda_i \in \{\emptyset, 1, 2\}$, $\emptyset$ represents an 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 $s = 2$

simplifications

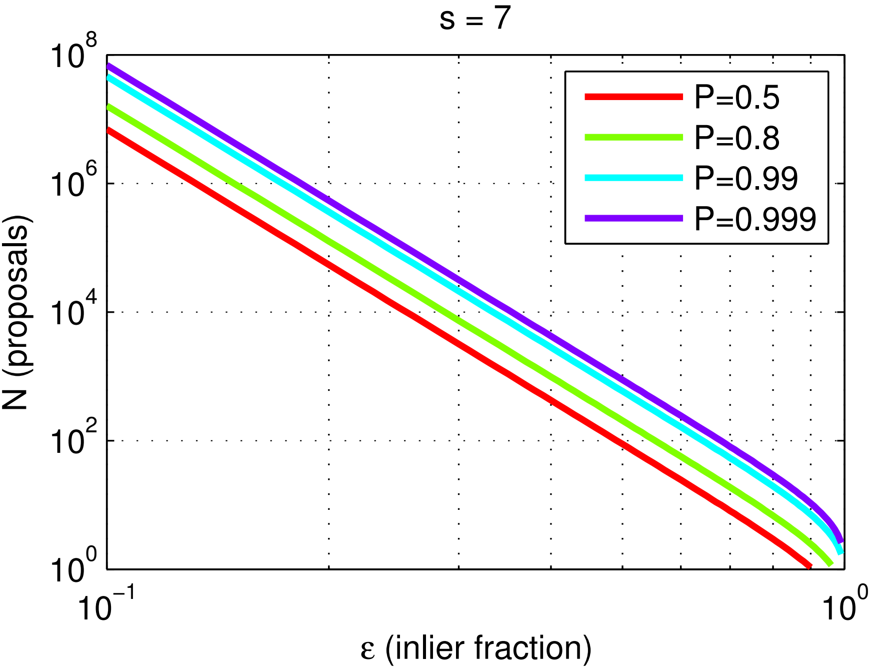
- vanishing points restricted to the set of all pairwise segment intersections
- mother lines fixed by segment centroid, then θ_L uniquely given by λ_i, and the configuration is

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



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

Thank You



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