Pointless Global Bundle Adjustment with Relative Motions Hessians

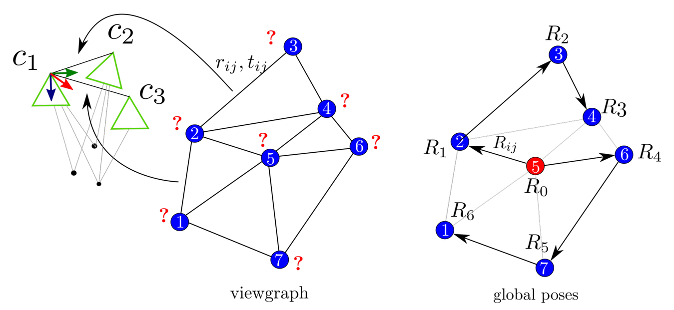
Ewelina Rupnik, Marc Pierrot Deseilligny
LaSTIG, UGE, IGN-ENSG erupnik.github.io/pointlessGBA.html

PCV WORKSHOP CVPR JUNE 18-22, 2023 VANCOUVER, CANADA

Camera pose estimation

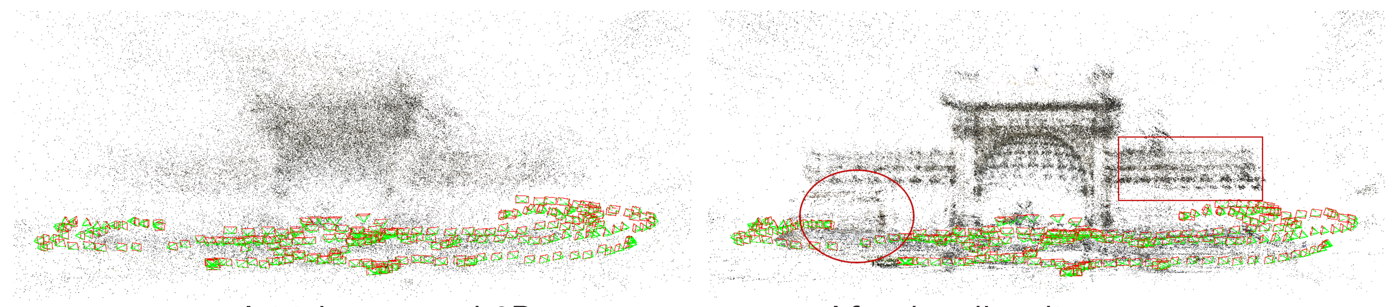
Our problem:

1. Camera pose initialisation:
   - Relative $\{r_{ij}, \mathbf{c_{ij}}\}$ → in parallel & fast
   - Global R_i, then $\mathbf{C_i}$ → fast but not precise



2. Refinement: bundle adjustment (BA) → rigorous & precise but slow

$$E_{BA} = \sum\sum \rho \cdot \mathbf{r}^2 = \sum_{i}^{N}\sum_{j}^{M} \rho \cdot [proj(\mathbf{p}_i) - \mathbf{o}_{ki}]^2 \ .$$



Initial poses and 3D — After bundle adjustment

Key ideas:

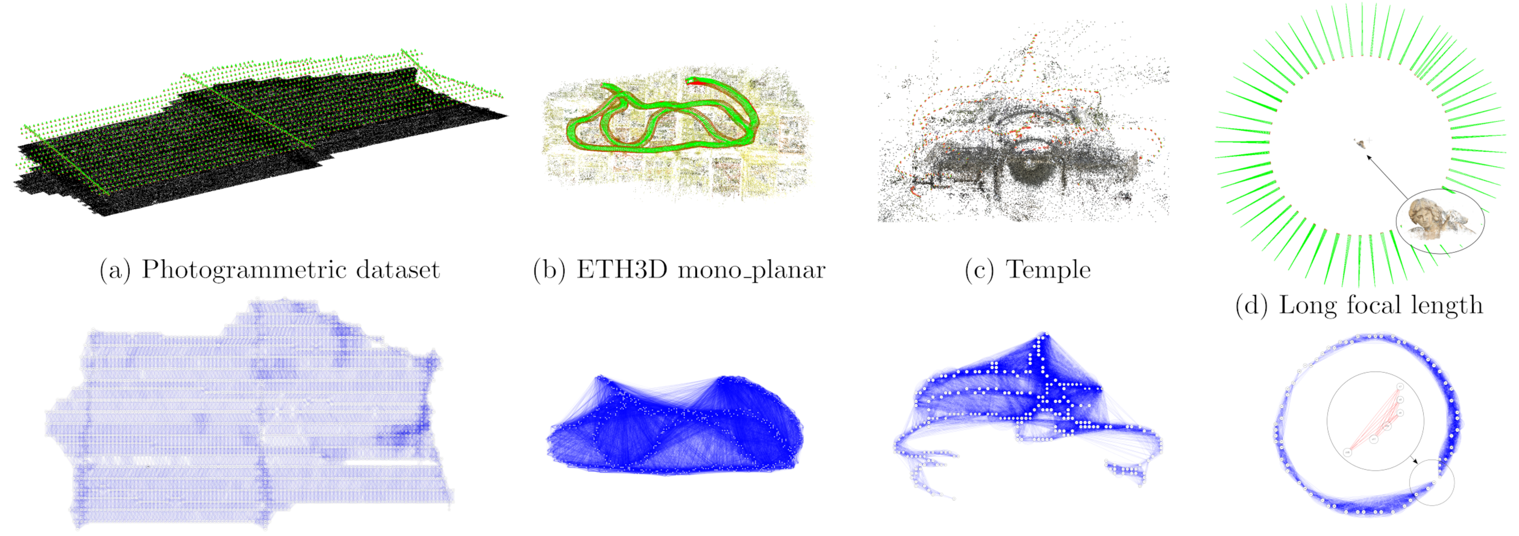
- ✅ Local BAs over relative motions provide **hessians**
- ✅ Hessians encode **features' random errors** & **correlations** between camera poses and 3D points
- ❌ Existing methods **discard** those information rich matrices

***Pointless* BA:**

- ⭐ **Leverage hessians** in global BA while **excluding features**
- ⭐ **Propagate** the locally defined hessians to the global frame & **adjust** → rigorous & precise & fast & low cost

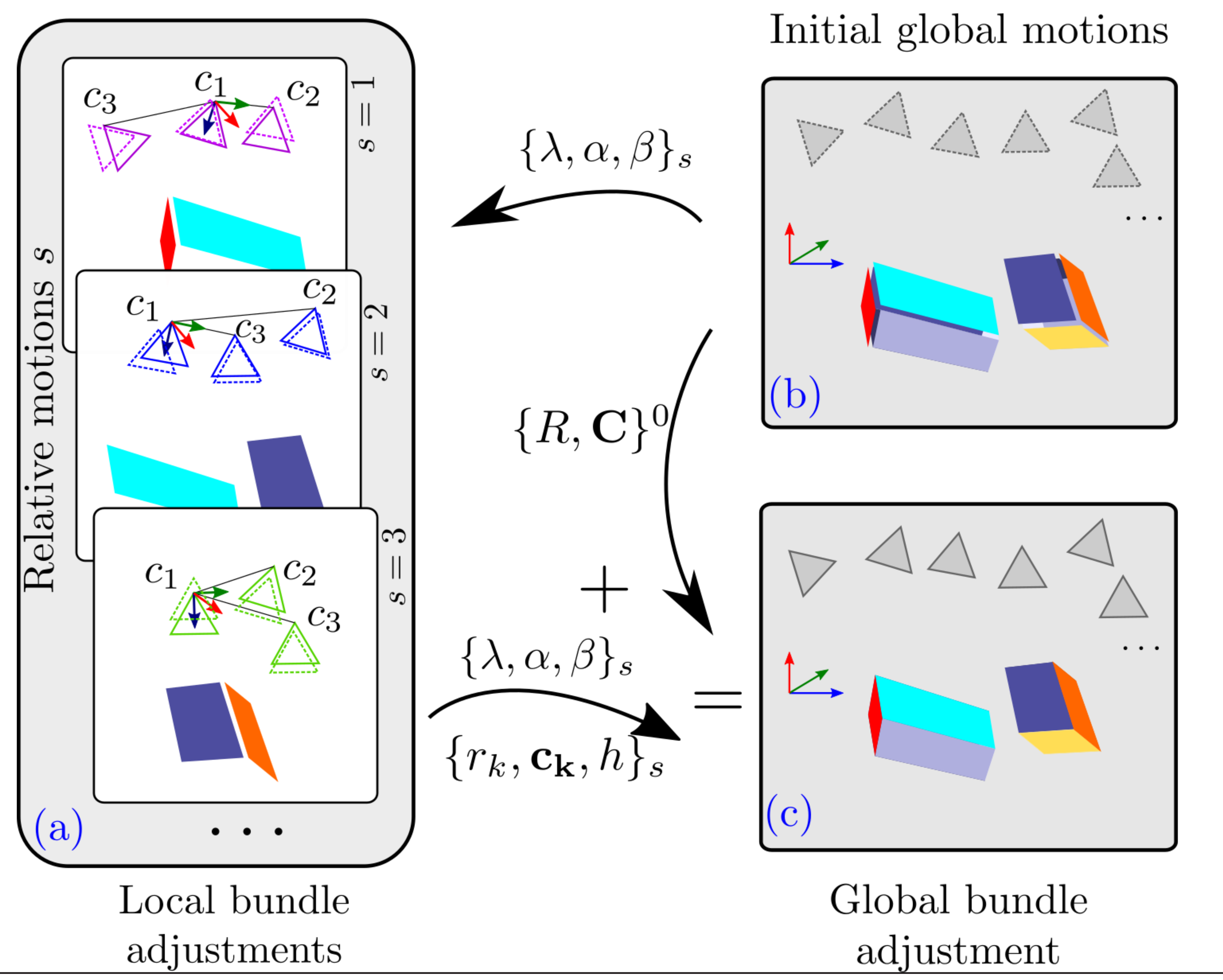
Datasets

×1 photogrammetric d., ×2 computer vision benchmark, ×1 long focal d.



(a) Photogrammetric dataset (b) ETH3D mono_planar (c) Temple (d) Long focal length

Method



1. Compute local motions & retrieve local Hessians w.r.t. camera parameters

$$E_{BA} = \sum\sum \rho \cdot \mathbf{r}^2 = \rho\left(\delta x^T \overbrace{(J^T J)}^{H} \delta x + r_0^T J \delta x + r_0^2\right)$$

2. Extract the *camera reduced matrix* h from H using Schur complement
3. Find initial global poses

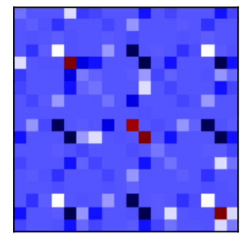
$$x_s = \{c, rot\} = \{\lambda_s \alpha_s C + \beta_s, \alpha_s R\}$$

4. Refine global poses

Averaging $\quad E^g_{BA} = \sum_s^S \sum_i^3 \rho \| \{c_{si}, rot_{si}\} - \{\lambda_s \alpha_s C + \beta_s, \alpha_s R\} \|^2$

Pointless BA $\quad E^s_{BA} = \delta x^T h \delta x + \cdots \quad \Rightarrow E^g_{BA} = \sum_{s=0}^S E^s_{BA}$

$\delta x = \{\lambda \alpha C + \beta, \alpha R\} - \{c_0, rot_0\}$



Local Hessians

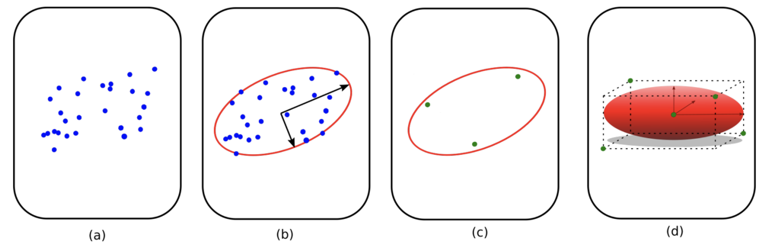
- Rigorous: encode stochastics of relative motions
- Compact: information stored in $6N \times 6N$ matrices
- Low cost: by-product of local BA (parallel comp.)

Global Bundle Adjustment

- Fast: disengages feature points from the adjustment
- Precise: propagates the *quality* of local BA to the global frame
- Flexible: can be easily integrated within any SfM pipeline

5-Points Bundle Adjustment

Idea: 1 Reduce features per motion 2 refine with classical BA.

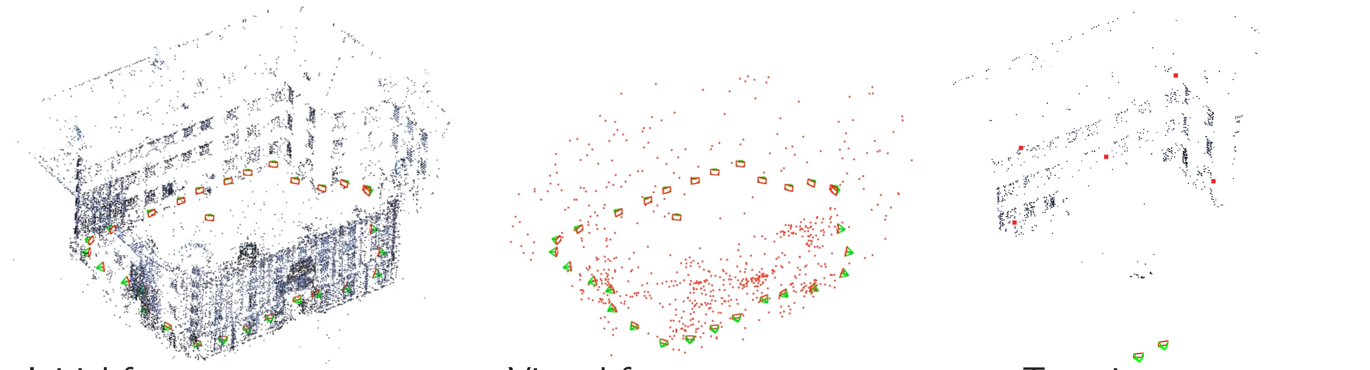


- Triangulate features
- Inscribe ellipsoid in 3D features
- Generate virtual features

$$\mu = \frac{1}{\sum_{i=1}^N P_i}\sum_{i=1}^{N} P_i \cdot Pds_{P_i}, \quad K_P = \frac{1}{\sum_{i=1}^N P_i}\sum_{i=1}^{N} Pds^2 \cdot P_i P_i^T - \mu_P \cdot \mu_P^T$$

$$\mathbf{P_i} = \mu + e_1 \mathbf{E_1} \cdot p_{i_1} + e_2 \mathbf{E_2} \cdot p_{i_2} + e_3 \mathbf{E_3} \cdot p_{i_3},$$

$$p_1 = [-1,1,1], p_2 = [1,-1,1], p_3 = [1,1,-1], p_4 = [0,0,0], p_5 = [-1,-1,-1]$$

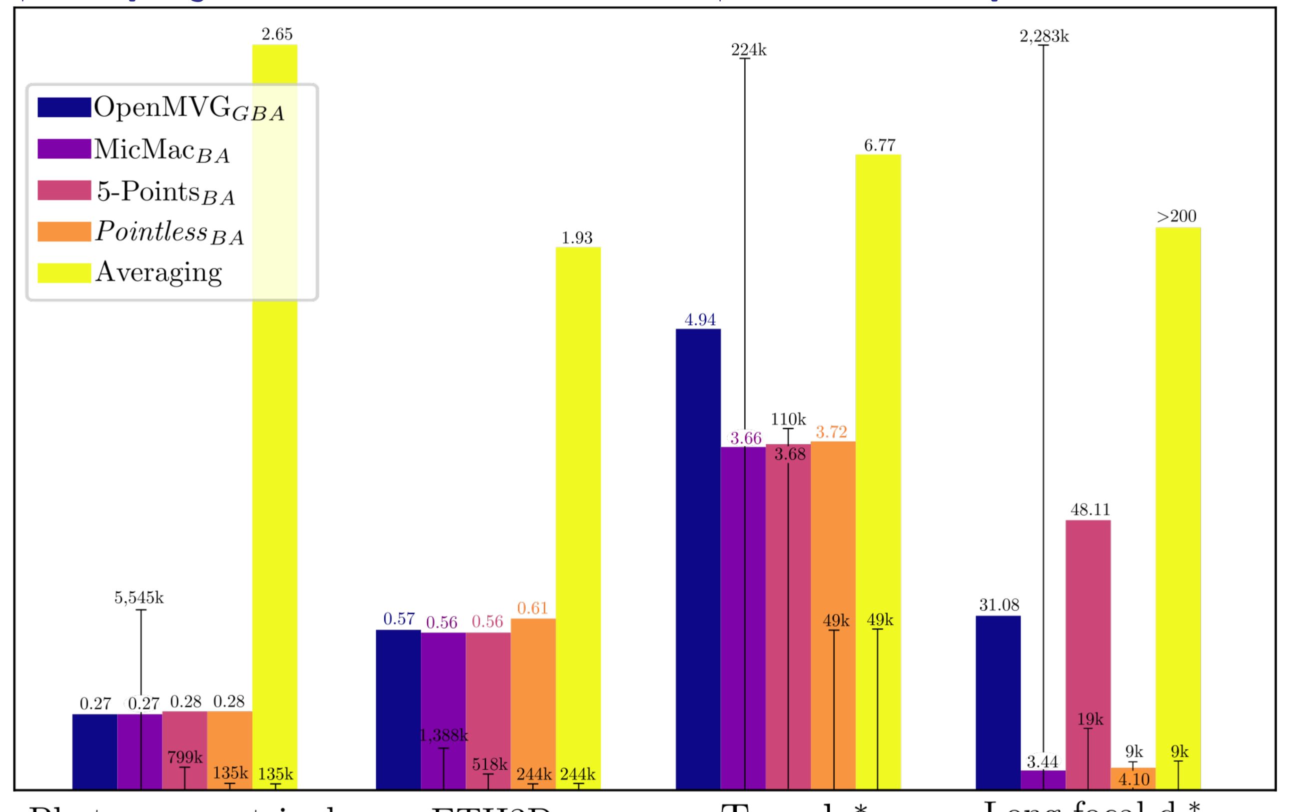


Initial features — Virtual features — Two-view case

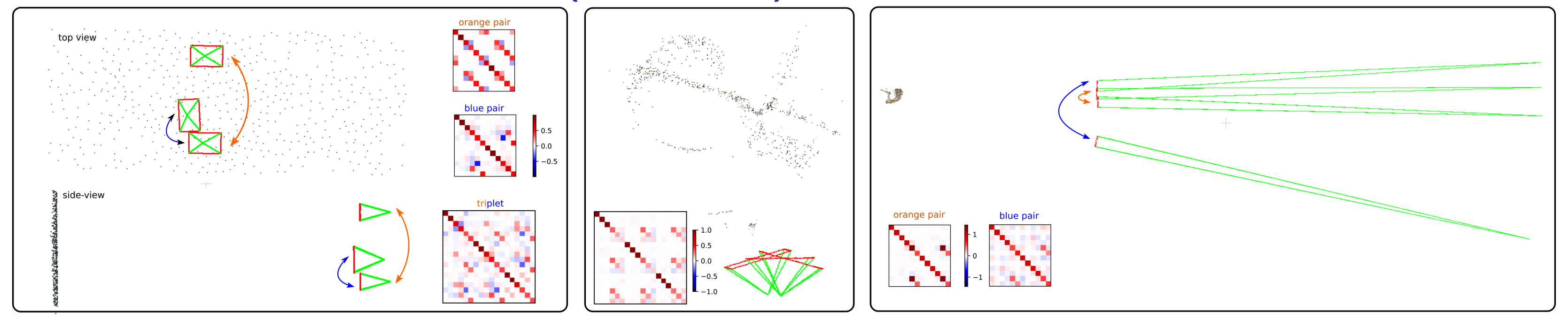
🟢🟡🟠	Fast	Rigorous	Precise	in existing SfM
5-Points BA:	🟢	🟠	🟡	🟢
***Pointless* BA:**	🟢	🟢	🟢	🟡

Experiments & Results

↓ Re-projection errors, ↓ number of parameters



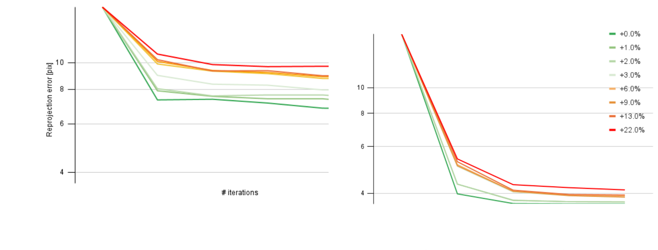
Covariance matrices (hessian^{-1}) vs geometry of relative motion



Initial 3D structure

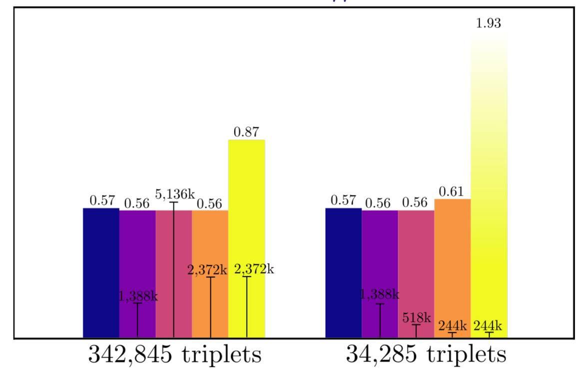
Averaging — 5-Points BA — *Pointless* BA — BA

Sensitivity to outliers ↓

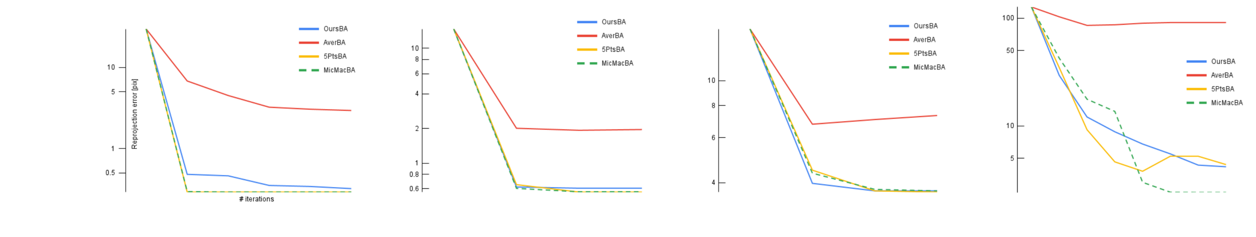


Averaging — *Pointless* BA

Performance vs # of motions



Convergence rates ↓



Photogrammetric d. — ETH3D — Temple — Long focal

🟢🟠	Averaging	*Pointless* BA	BA
Excludes features	🟢	🟢	🟠
Includes random feature errors	🟠	🟢	🟢
Efficiency & #overlapping motions	🟠	🟠	∅
Considers 📷 extrinsics' uncertainties & correlations related to geometry	🟠	🟢	🟠