BAFS: Bundle Adjustment with Feature Scale Constraints for Enhanced Estimation Accuracy

[1]Technion - Israel Institute of Technology, Israel

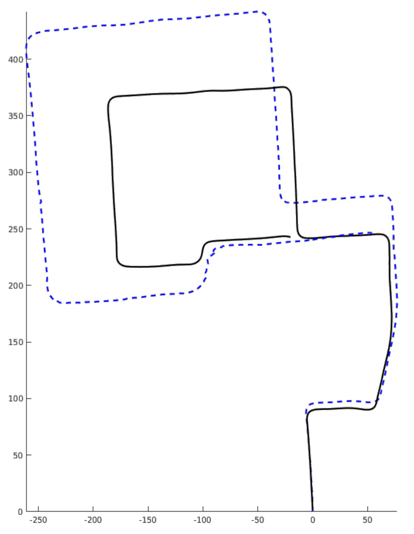
Vladimir Ovechkin[1] and Vadim Indelman[1]

ANPL Autonomous Navigation and Perception Lab

TECHNION Israel Institute of Technology

1. Introduction

- **Scale drift** is a known problem in monocular SLAM
- Without assuming prior information, camera motion and 3D map can be only estimated up to scale, which drifts over time
- Existing (non-learning) approaches either require loop closures, exploit non-holonomoic constraints or fuse information from other sensors



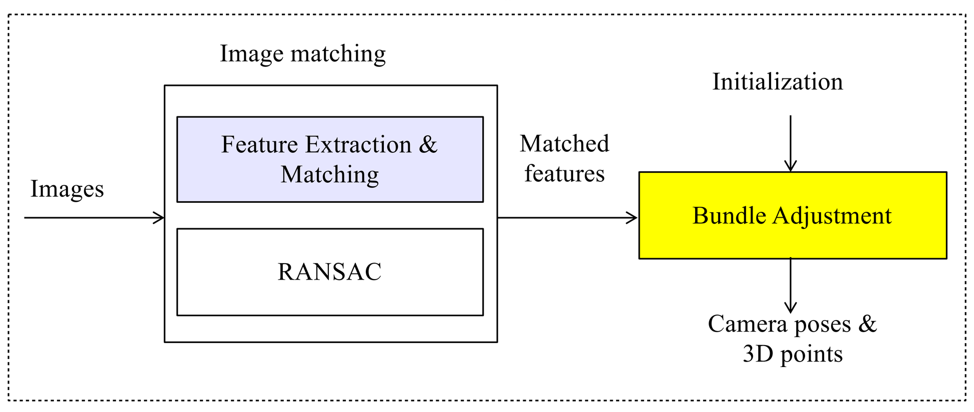
2. Contribution

- We formulate novel image feature scale constraints and incorporate these within bundle adjustment (BA)
- Leads to significantly improved estimation accuracy (especially) along the optical axis of the camera in a monocular setup
- Method does not require loop closures

3. Concept

- **Key observation**: detected feature scale changes consistently across a sequence of images
- Concept leverages the scale invariance property of SIFT detectors
- Detected feature scale can be predicted as a function of camera pose, landmark 3D coordinates and the corresponding 3D environment patch
- To provide more accurate detected feature scales, we increase the resolution of Gaussian kernels within the SIFT detector
- Feature scales are already typically calculated by common feature detectors (e.g., SIFT), thus far - used **only** for image matching
- We propose to exploit this available scale information for improving the performance of BA

Overall scheme of working framework



Thus far: scale information is used only for matching

<u>Our method</u>: incorporates scale information also into BA

4. Scale Constraint

- Denote by S_j the corresponding environment patch, or virtual size, centered around landmark l_j
- Scale observation model:

$$\underbrace{s_i^j}_{\text{measured}} = \underbrace{f\frac{S_j}{d_i^j}}_{\text{predicted}} + v_i \qquad v_i \sim N(0, \Sigma_{fs})$$

- d_i^j is the distance <u>along optical axis</u> from the camera pose x_i to landmark l_j

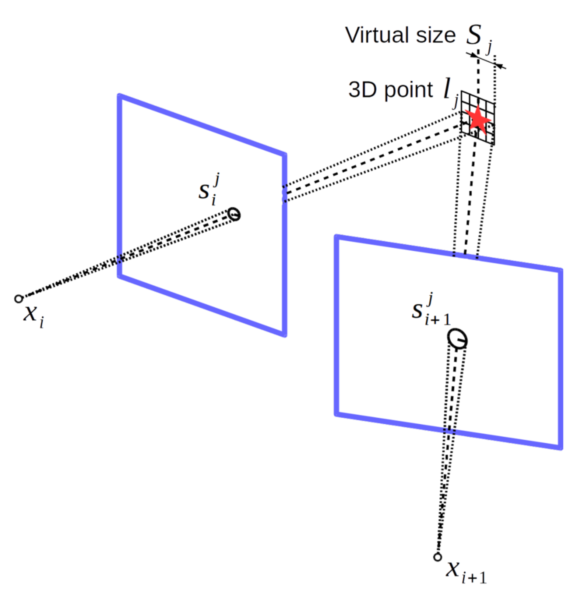
$$d_i^j(x_i, l_j) \doteq z_c, \quad \begin{bmatrix} x_c \\ y_c \\ z_c \end{bmatrix} = R_i l_j + t_i$$

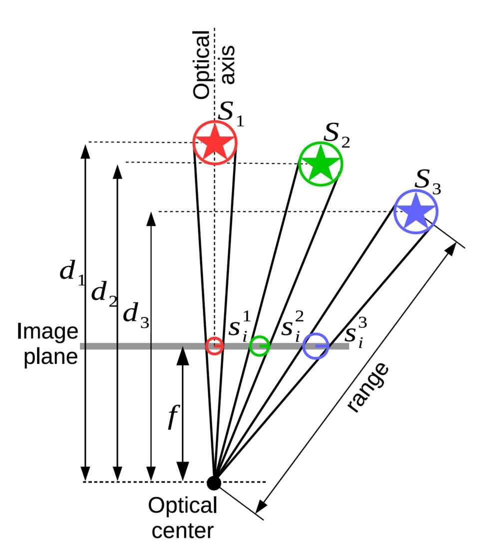
- Note d_i^j is not the range. To see why:

Virtual landmark size: $S_1 = S_2 = S_3$

Feature scales: $s_i^1 < s_i^2 < s_i^3$

Distance along optical axis: $d_1 > d_2 > d_3$





5. Incorporating Scale Constraint within BA

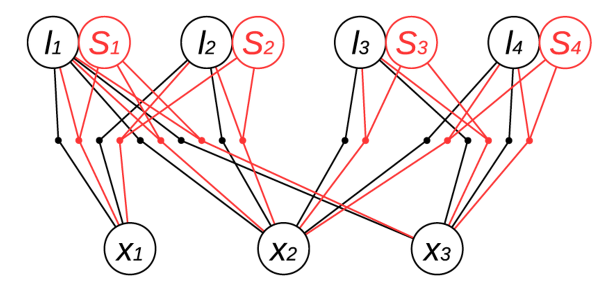
- Scale measurement likelihood: $\mathbb{P}(s_i^j|S_j,x_i,l_j) \doteq \frac{1}{\sqrt{|2\pi\Sigma_{fs}|}}\exp\left[-\frac{1}{2}\left\|s_i^j - f\frac{S_j}{d_i^j}\right\|^2_{\Sigma_{fs}}\right]$
- Joint posterior pdf: $\mathbb{P}(X,L,S|\mathcal{Z}) \propto \prod_i^N \prod_{j\in M_i} \mathbb{P}\left(z_i^j|x_i,l_j\right)\mathbb{P}\left(s_i^j|S_j,x_i,l_j\right)$
- Corresponding NLS problem: $J(X,L,S) = \sum_i^N \sum_{j\in M_i} \underbrace{\left\|z_i^j - \pi(x_i,l_j)\right\|^2_{\Sigma_v}}_{\text{re-projection error}} + \underbrace{\left\|s_i^j - f\frac{S_j}{d_i^j}\right\|^2_{\Sigma_{fs}}}_{\text{scale error}}$
- Initialization of virtual landmark size variable S_j:

  Triangulate landmark l_j

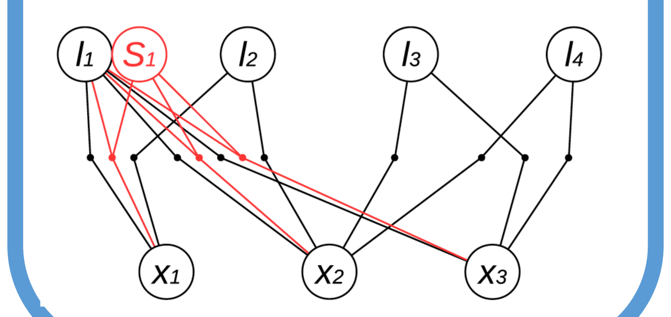
  Use current estimates of camera pose and landmark coordinates to calculate $S_j = s_i^j \frac{d_i^j}{f}$

6. Factor Graph

Naïve: all scale constraints

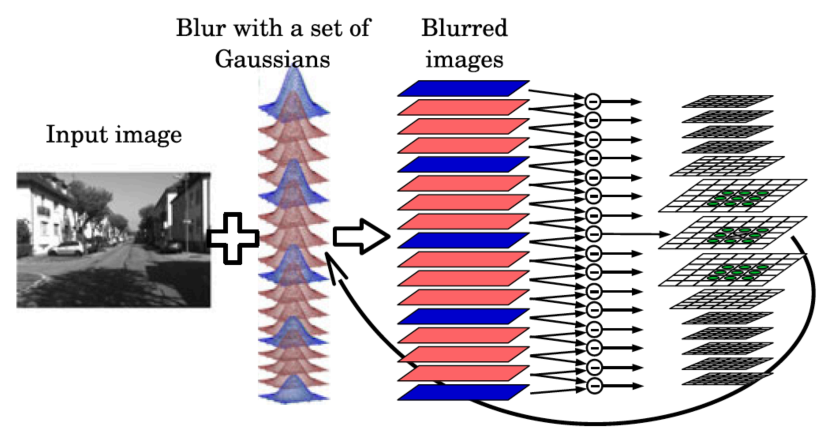


Heuristic: scale constraints only for long-track features



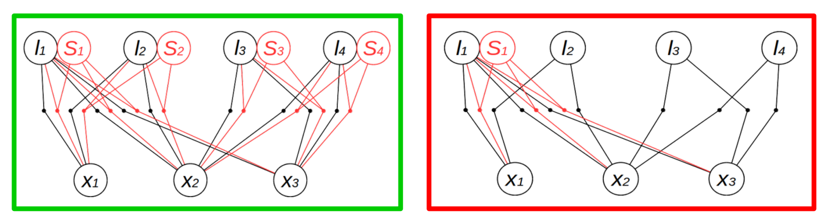
7. Enhancement of Feature Scale Measurement Accuracy

- Estimation accuracy is improved only if feature scale measurements are sufficiently accurate
- Simple method: get higher-accuracy scale measurements by increasing number of layers per octave (finer resolution of the Gaussian kernels)



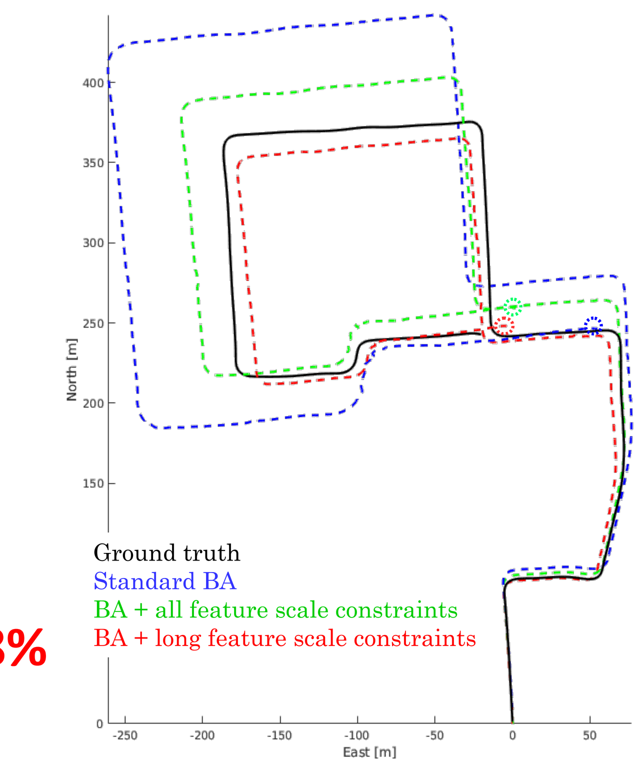
8. Results – KITTI Dataset

Difference between **red** and **green** track: amount of scale constraints.

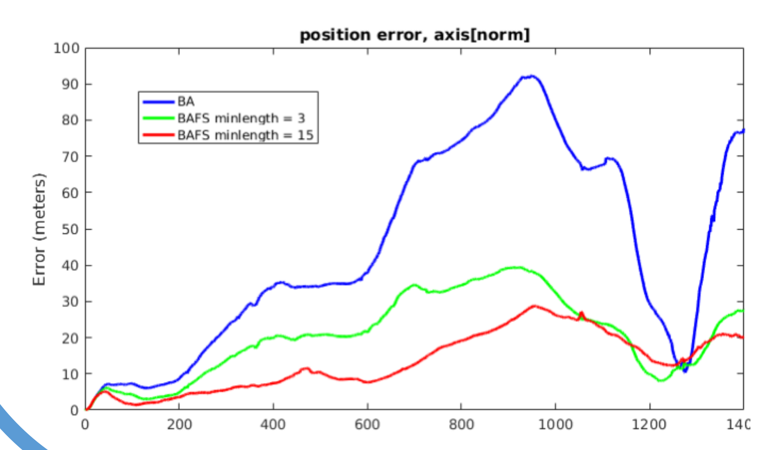


Error rate along optical axis:

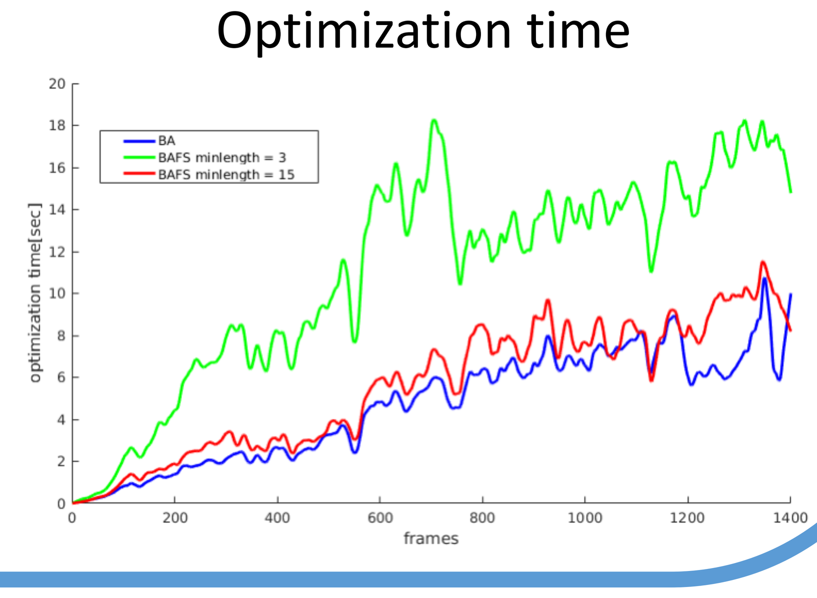
- first 100 frames: **8.6%** | **5%** | **1%**
- Last 100 frames: **94.6%** | **27.3%** | **8.8%**



Position error



Optimization time



9. Conclusions

Improved accuracy of bundle adjustment and monocular SLAM along optical axis direction, without requiring loop closures:

- Developed and introduced scale constraints within BA
- Feature scale information is already available from feature detector
- Enhanced feature scale measurement

RA-L with ICRA option (RA-L) WeA 2190