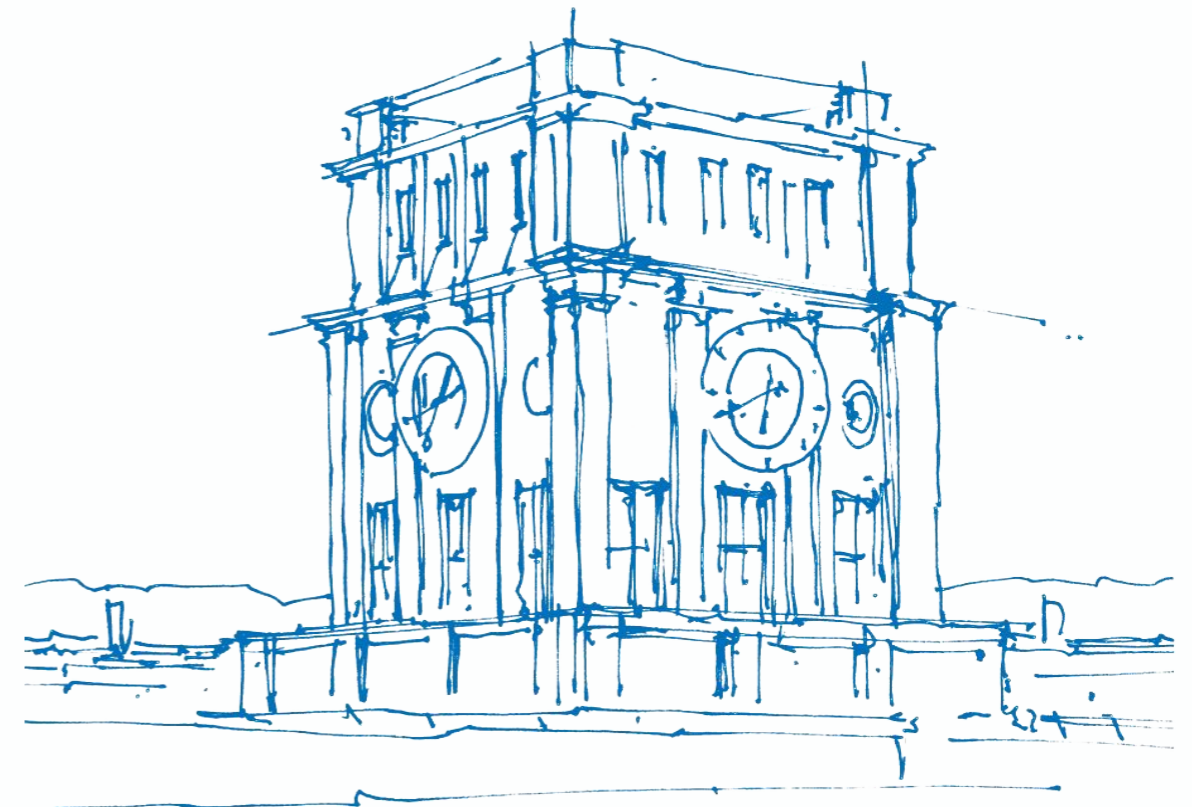


Practical Course: Vision Based Navigation

Projects

Jason Chui, Simon Klenk, Sergei Solonets
Prof. Dr. Daniel Cremers

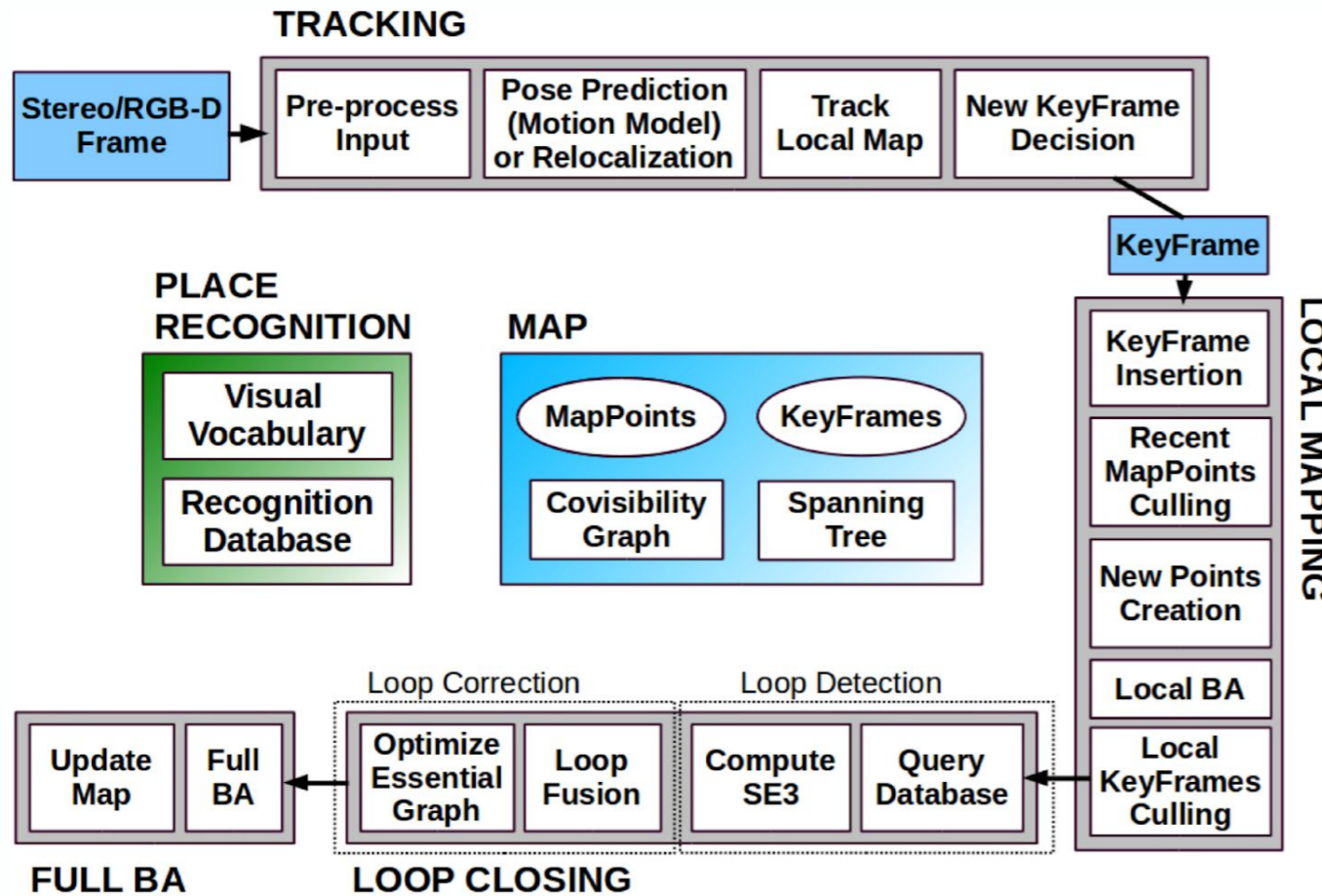


TUM Uhrenturm

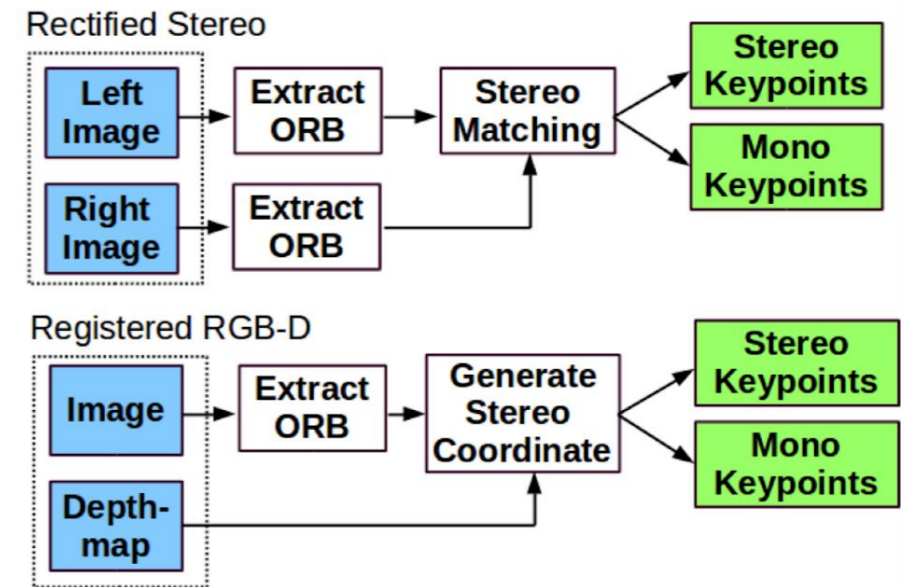
Project Phase

- Start after sheet 5 is complete. For the remainder of the lecture period.
- Work on open-ended project (1-2 people).
- Register your group and select your project topic and meeting time slot [here](#) by 06.12.2022
- Assign your group a tutor before 12.12.2022
- Mandatory weekly meeting with tutors to discuss progress and next steps.
 - fixed 30 min time slot
 - Mondays between 2pm-6pm
- Project goal is to be determined:
 - Choose from list of suggested projects or suggest your own.
 - “Advanced” topics have more uncertain scope / solution. More independent work required.
 - At most 2 groups should work on the same project (first come first served).
- Present project outcome in talk and Q&A session (15min + 5min)
- Written report on project outcome (10-12 pages, single column, single-spaced lines, 11 pt)
- Important dates:
 - Fix groups, project topic, and time for weekly meeting: 06.12.2022
 - First meeting on 12.12
 - Other meeting date on 19.12, 09.01, 16.01, 23.01
 - Project presentations: 31.01.2023 2:00pm
 - Project report due: 15.02.2023

1. SLAM



(a) System Threads and Modules.

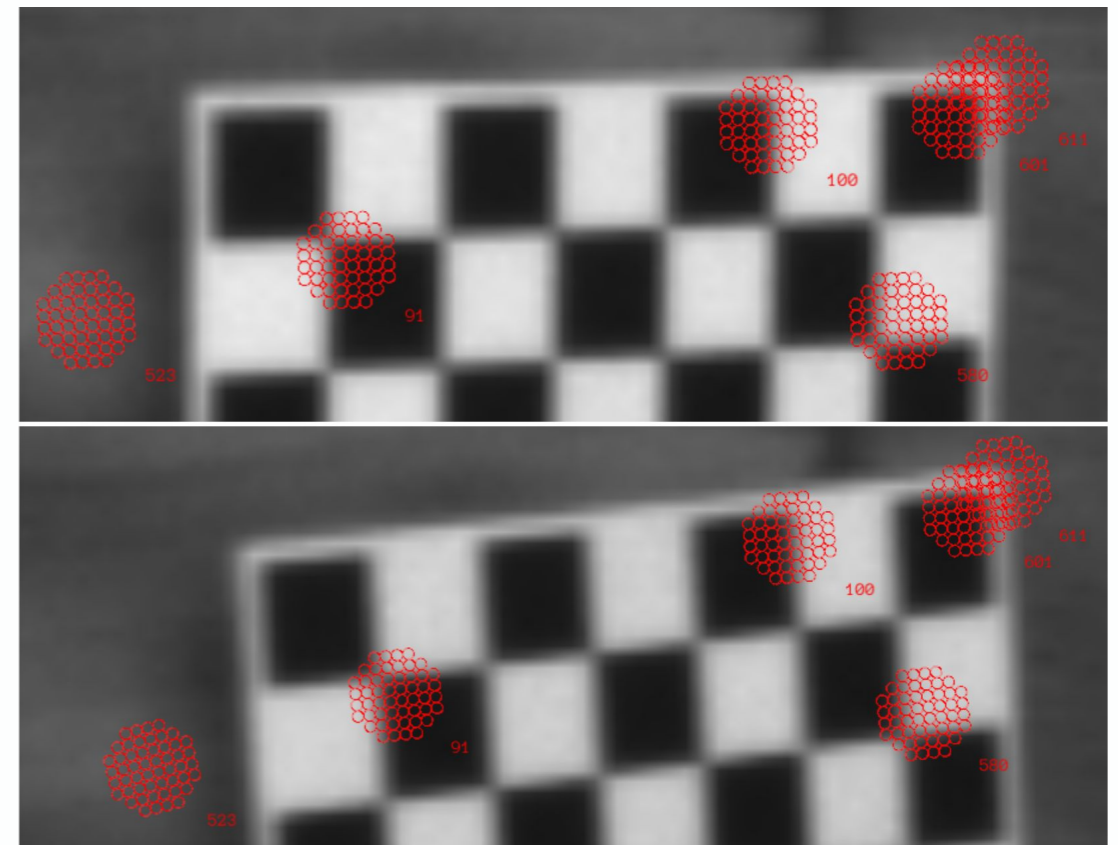


(b) Input pre-processing

- ORB_SLAM: <http://webdiis.unizar.es/~raulmur/MurMontielTardosTRO15.pdf>
- ORB_SLAM2: <https://arxiv.org/abs/1610.06475>
- Map management
- Reusing Keyframes
- Spanning tree for pose-graph optimization

2. Indirect Visual Odometry with Optical Flow

- Sparse optical flow as alternative to feature matching
- Extend odometry application
- Compare runtime, accuracy, ...
- Possible extensions:
 - patch similarity norms
 - Keyframing, local optimization
 - Different image warping strategies
 - Implement Gauss-Newton (or LM) manually

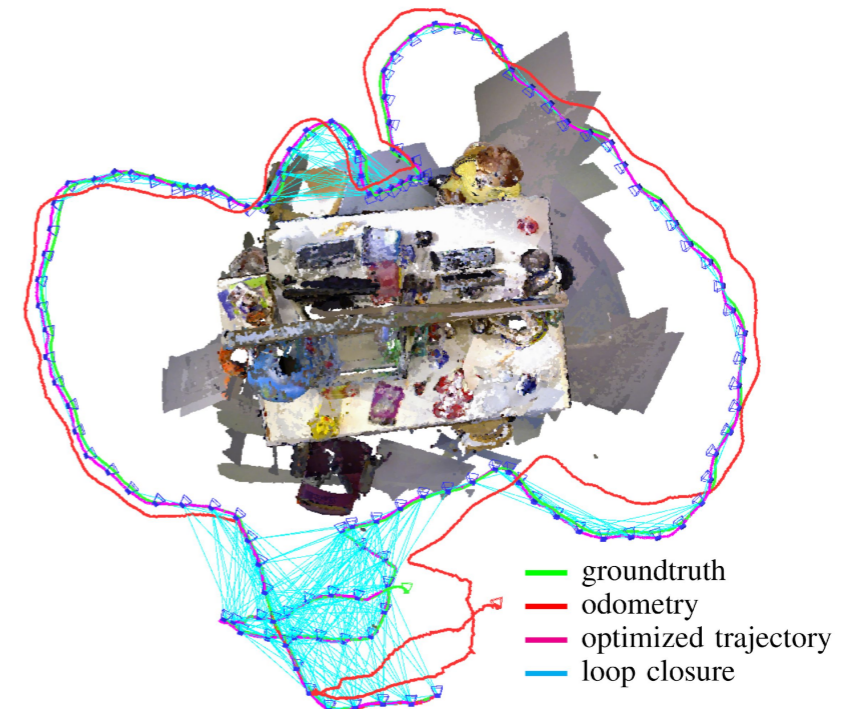
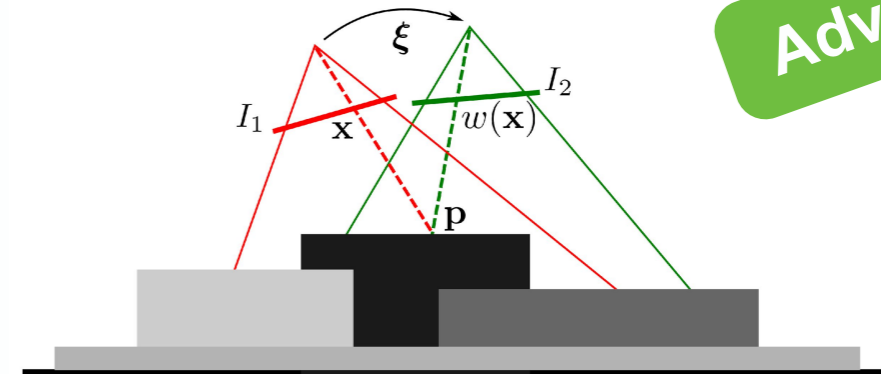


- Visual-Inertial Mapping with Non-Linear Factor Recovery (V. Usenko, N. Demmel, D. Schubert, J. Stueckler and D. Cremers), In arXiv:1904.06504, 2019. <https://arxiv.org/pdf/1904.06504>
- Equivalence and efficiency of image alignment algorithms (Baker, Simon, and Iain Matthews), In IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 1. IEEE Computer Society; 1999, 2001. <http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.70.20&rep=rep1&type=pdf>

3. Direct Visual Odometry for RGB-D Images

Advanced

- Work with RGB-D data
- Estimate the relative pose via Direct Image Alignment
- Implement Gauss-Newton (or LM) manually
- Frame-to-frame or frame-to-keyframe
- Different image warping strategies
- coarse-to-fine to improve convergence
- robust-norm to handle outliers



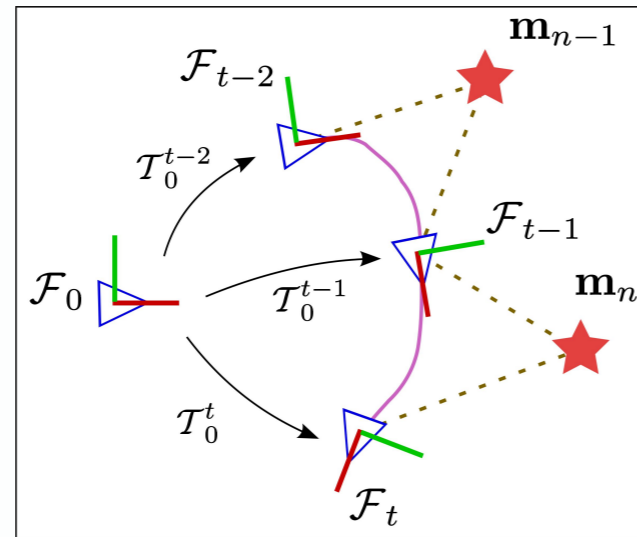
- Robust Odometry Estimation for RGB-D Cameras (C. Kerl, J. Sturm and D. Cremers), In International Conference on Robotics and Automation (ICRA), 2013.

<https://vision.in.tum.de/media/spezial/bib/kerl13icra.pdf>

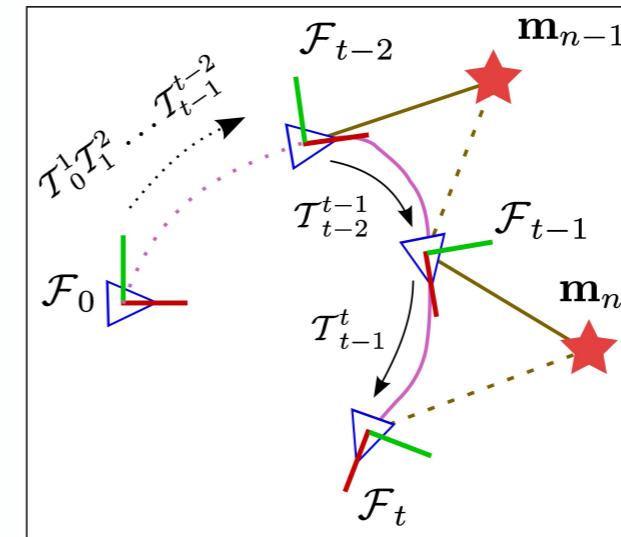
- Equivalence and efficiency of image alignment algorithms (Baker, Simon, and Iain Matthews), In IEEE Computer Society Conference on Computer Vision and Pattern Recognition. Vol. 1. IEEE Computer Society; 1999, 2001.

<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.70.20&rep=rep1&type=pdf>

4. Relative Map Formulation for SLAM

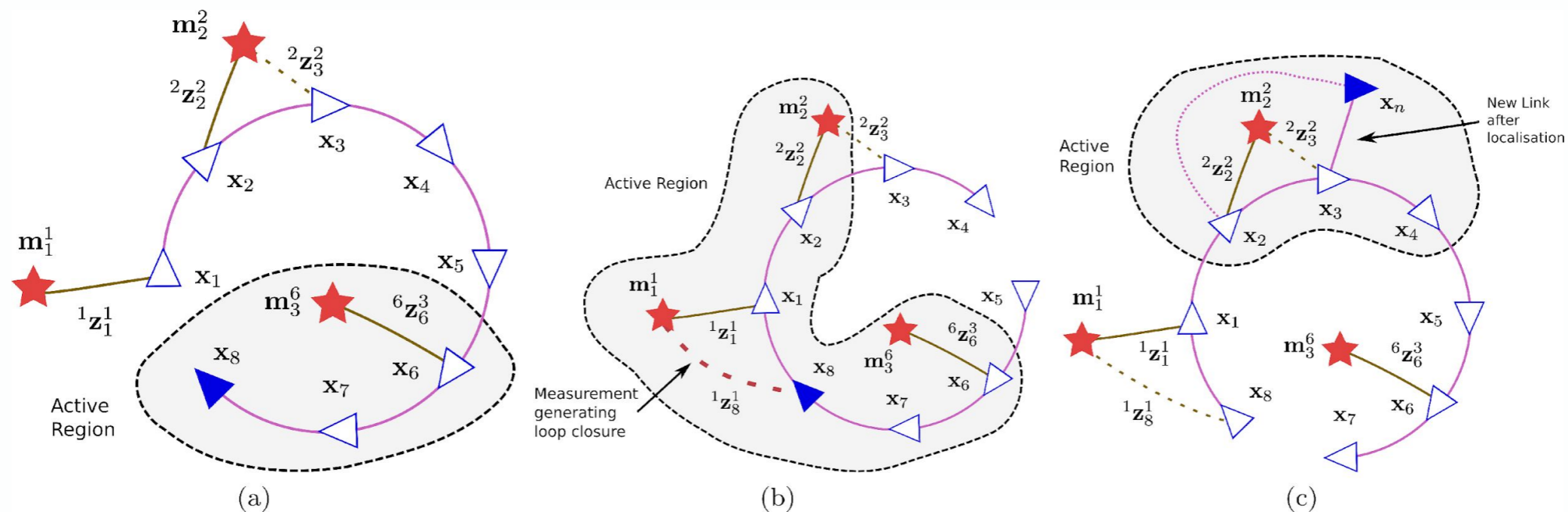


(a) Global

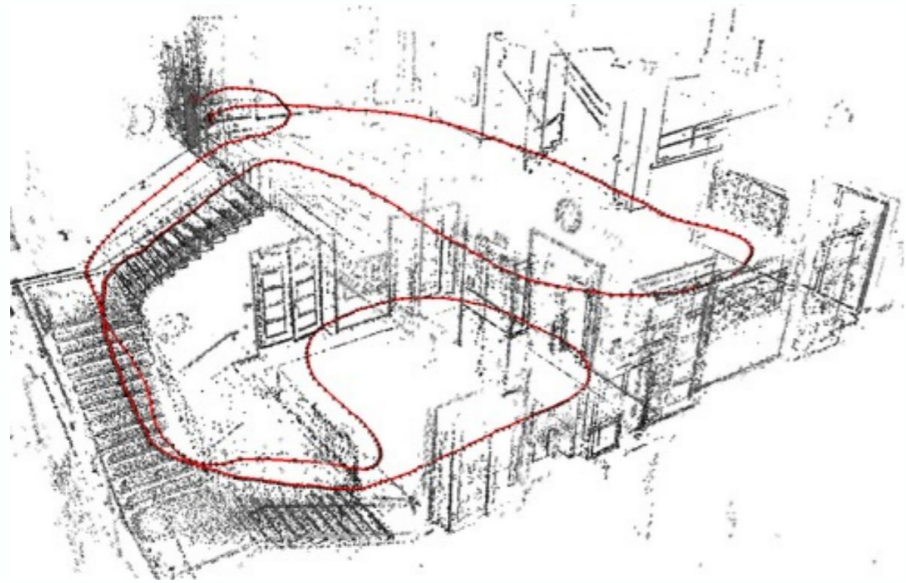


(e) Continuous relative representation (CRR)

- Change the map formulation to the relative one
 - Parameters are relative poses between keyframes
 - All points are defined relative to some frame
- Extend either SfM or Odometry application
- Paper: http://www.robots.ox.ac.uk/~mobile/Papers/2010IJCV_mei.pdf



5. Photometric Bundle Adjustment

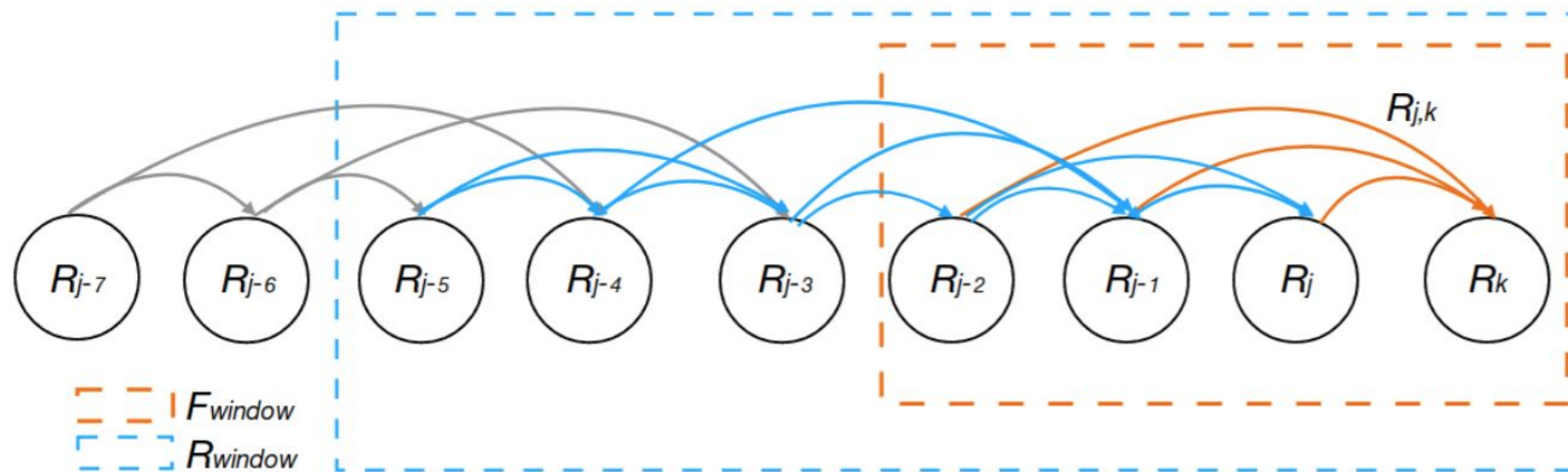


$$E_{\mathbf{p}j} := \sum_{\mathbf{p} \in \mathcal{N}_{\mathbf{p}}} w_{\mathbf{p}} \left\| \left(I_j[\mathbf{p}'] - b_j \right) - \frac{t_j e^{a_j}}{t_i e^{a_i}} \left(I_i[\mathbf{p}] - b_i \right) \right\|_{\gamma}$$

- Photometric Bundle adjustment in SFM
 - Error metric similar to DSO (<https://arxiv.org/pdf/1607.02565.pdf>)
 - Initialize and optimize additional (non-feature) points
 - Possibly use vignetting and response
- Very advanced: extend a recent DSO implementation (<https://github.com/RoadlyInc/DSOPP>) to stereo, or change its map representation. This project does not rely on your current code framework, and hence requires significant additional effort.

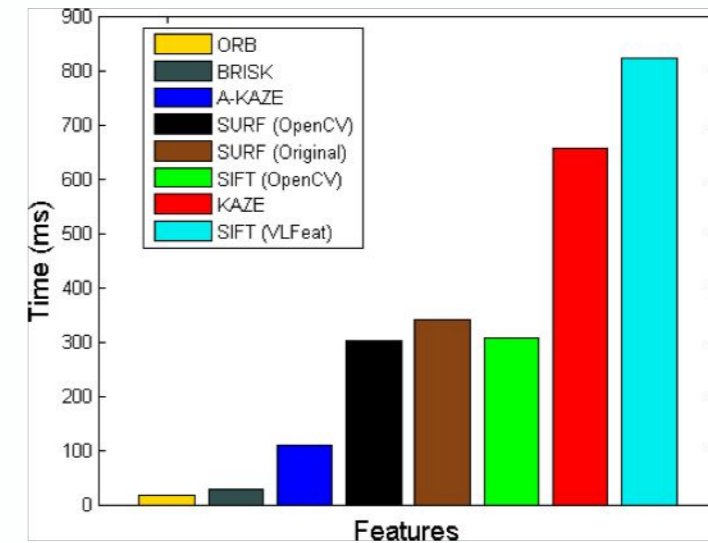
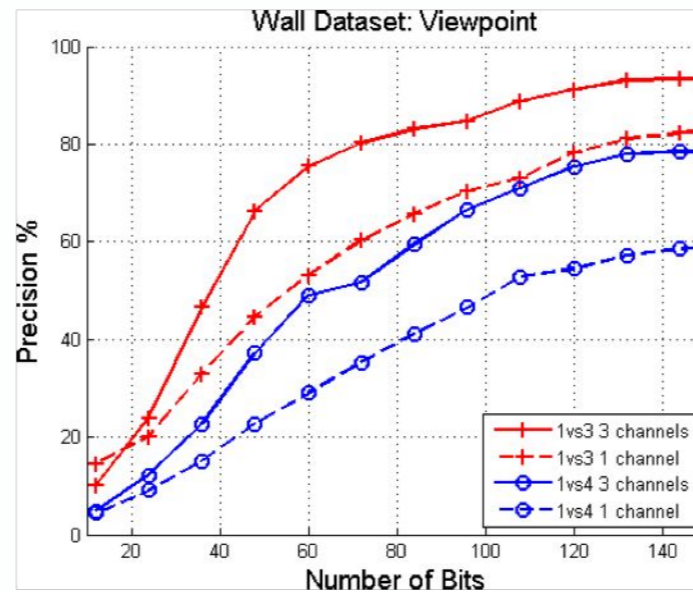
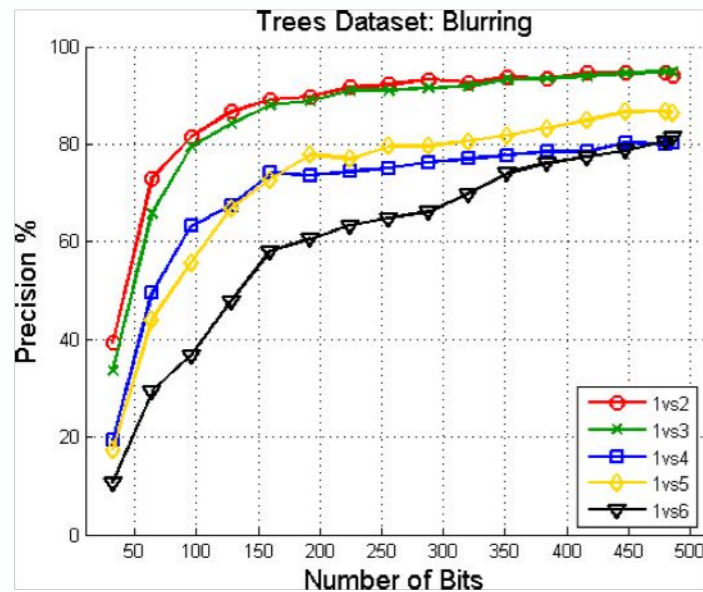
6. Rotation-only SLAM

Advanced



- Chng et al, “Monocular Rotational Odometry with Incremental Rotation Averaging and Loop Closure”, 2020
- paper: <https://arxiv.org/pdf/2010.01872.pdf>
- Incremental rotation averaging
- Global optimization with “pose graph optimization” (but only rotations)
- Handle pure-rotation case for monocular camera
- Once rotations are estimated: “Known rotation” SLAM or SfM
- Possible extensions: S. LeeHun and J. Civera. “Rotation-Only Bundle Adjustment”, 2020 (<https://arxiv.org/pdf/2011.11724.pdf>)

7. Advanced Matching and Keypoint Evaluation

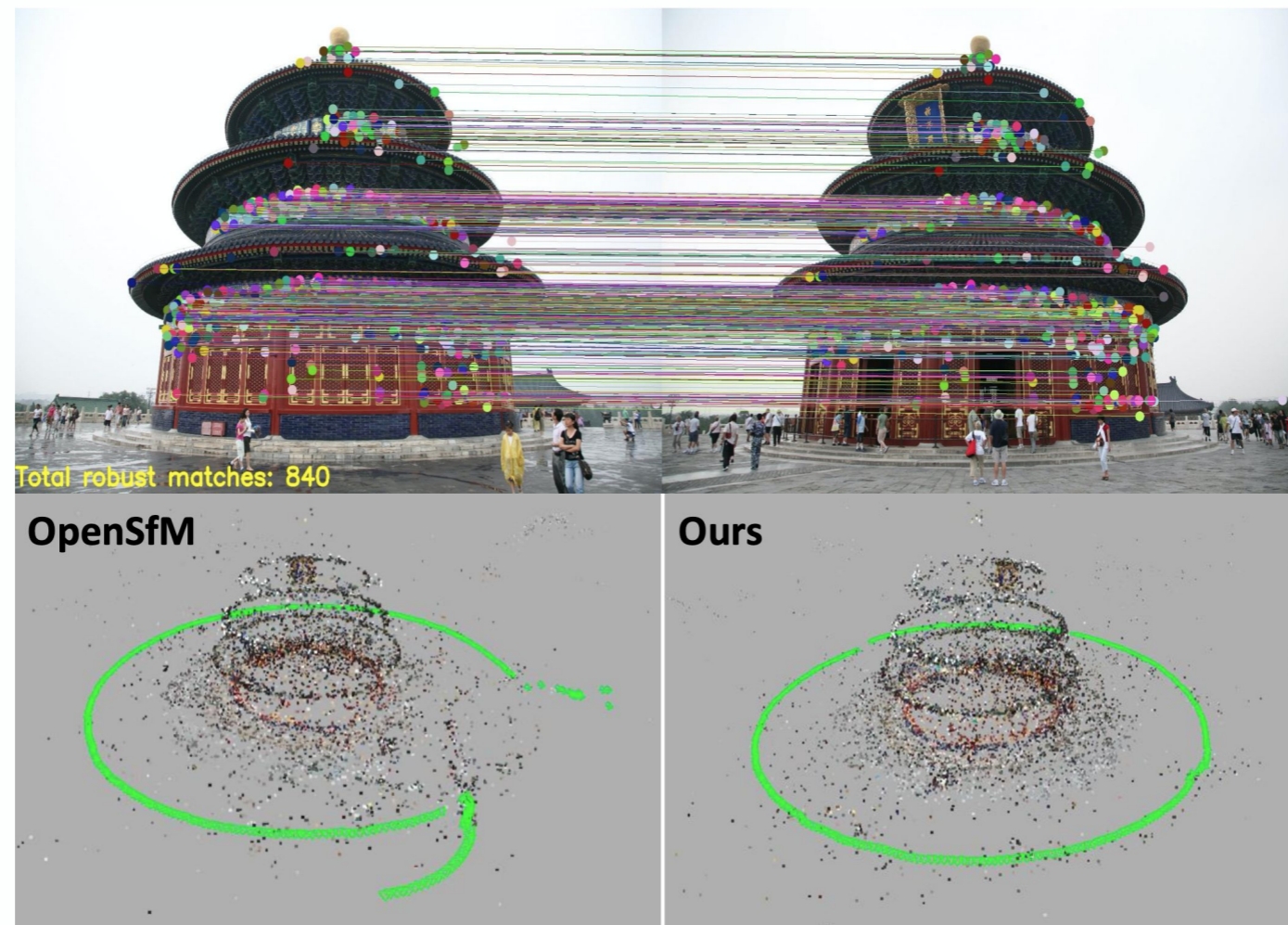


Alcantarilla, Pablo F., and T. Solutions. "Fast explicit diffusion for accelerated features in nonlinear scale spaces." *IEEE Trans. Patt. Anal. Mach. Intell* 34.7 (2011): 1281-1298.

- Keypoints evaluation:
 - ORB, AKAZE, SIFT, BRISK
 - Computation time / matching statistics
- Cascade Hashing for descriptor matching:
 - <http://www.nlpr.ia.ac.cn/jcheng/papers/CameraReady-CasHash.pdf>

8. Improving SfM with Reliable Resectioning

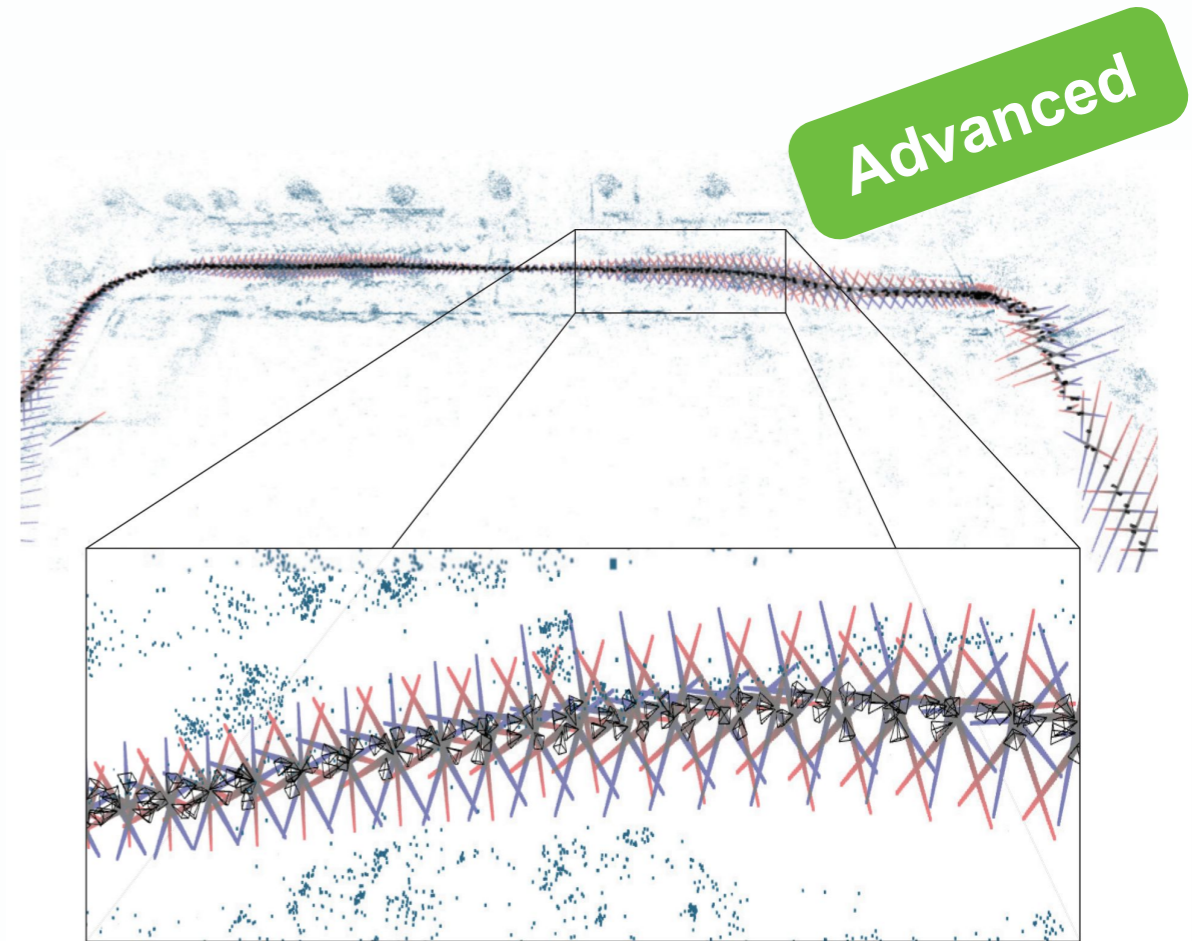
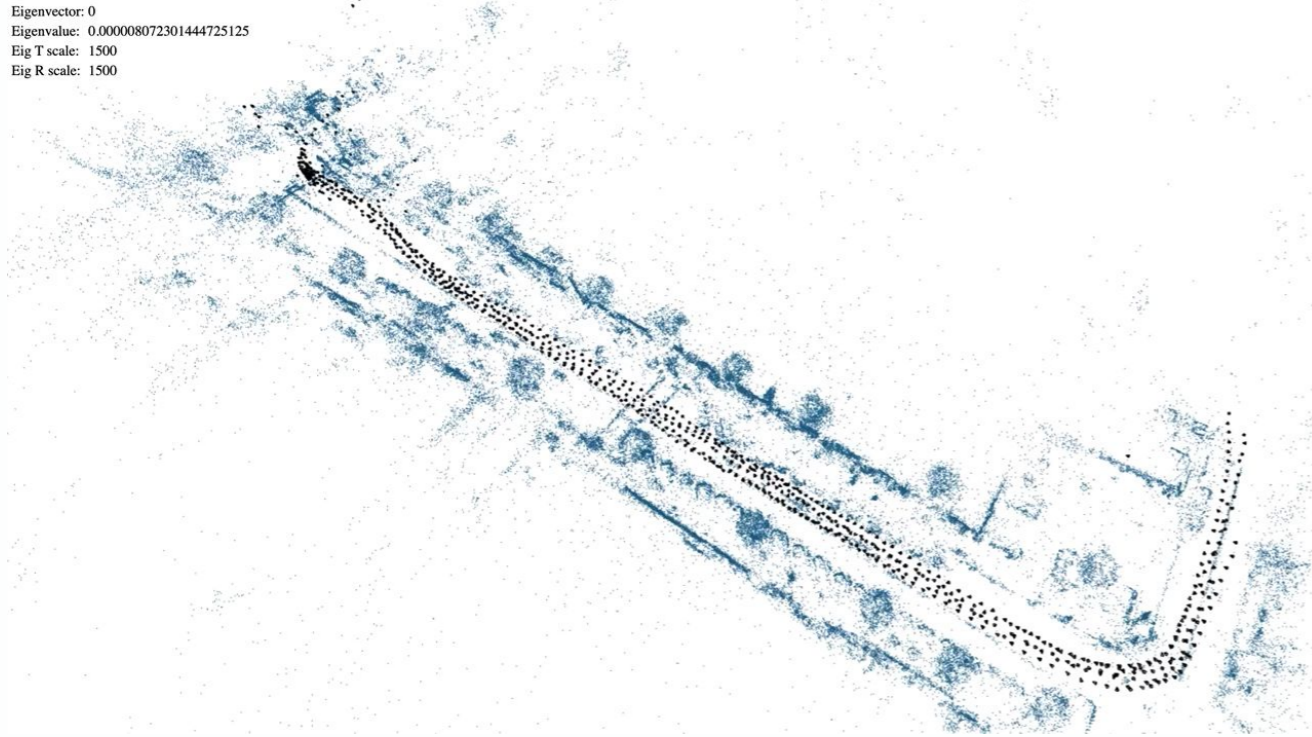
Advanced



- R. Kataria et al., 3DV 2020
 - paper & code: <https://github.com/rajkataria/ReliableResectioning>
 - video: <https://slideslive.com/embed/presentation/38941065>
- Implement ideas from paper in our SfM pipeline:
 - Adjust order of adding cameras (weight shorter tracks higher)
 - Initialise pose only with reliable matches; then find more matches by projection

9. Visualizing Spectral BA Uncertainty

Eigenvector: 0
Eigenvalue: 0.000008072301444725125
Eig T scale: 1500
Eig R scale: 1500

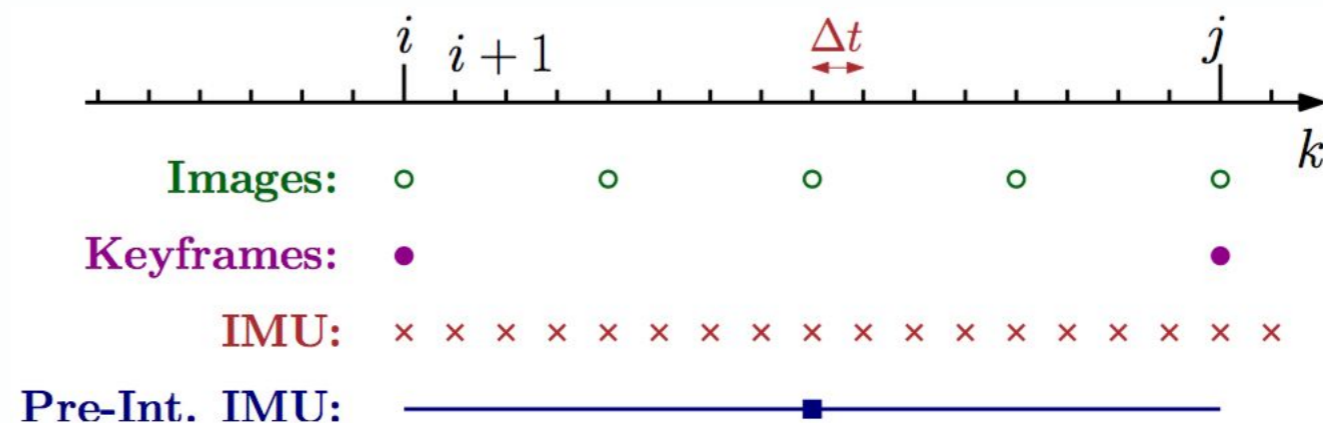
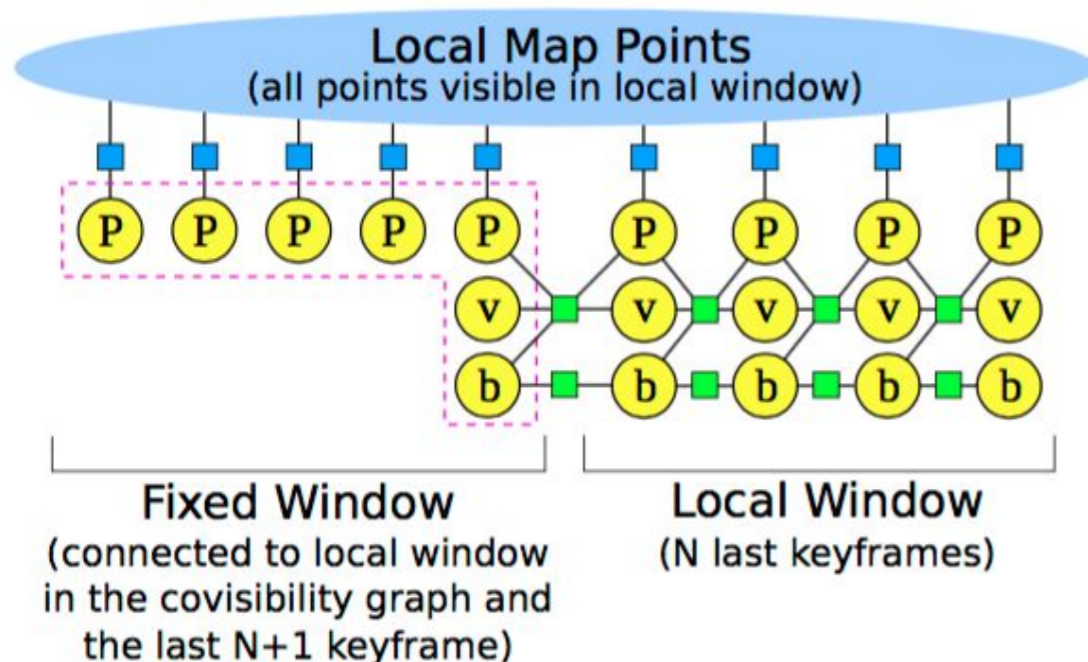


- K. Wilson and S. Wehrwein, 3DV 2020
- <https://facultyweb.cs.wvu.edu/~wehrwes/files/sfmflex.pdf>
- video: <https://slideslive.com/embed/presentation/38941084>
- Implement pose uncertainty visualisation in our SfM application
 - static (ellipses)
 - animated
- Investigate relative scaling of units (rotation vs translation)
- public implementation for reference: <https://wilsonkl.github.io/sfmflex-release/>

10. Visual-Inertial Tracking using Preintegrated Factors

Advanced

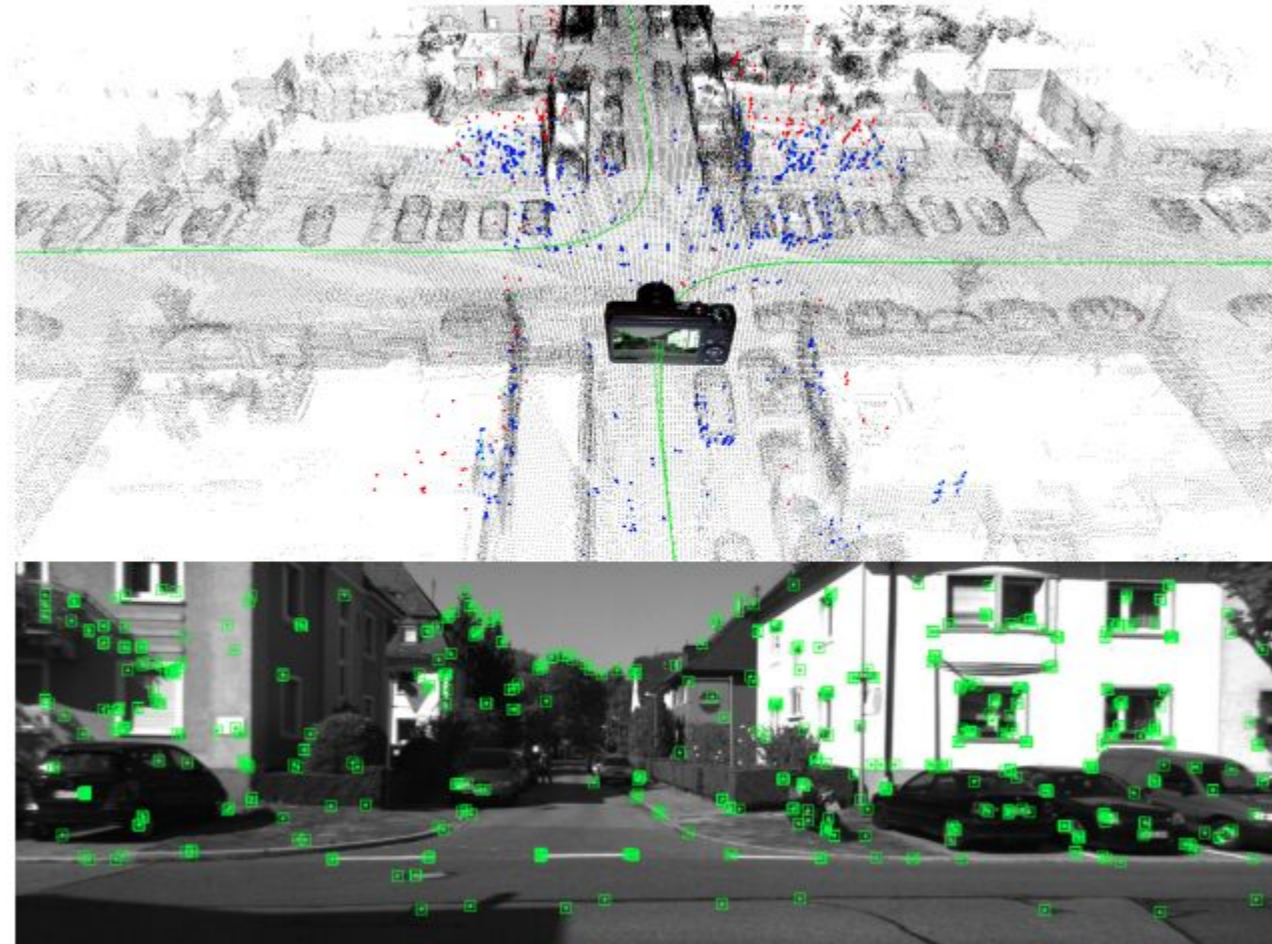
- Use camera + IMU for stability and scale observability
- Estimate IMU biases and velocity
- Preintegrate measurements between image frames



- Theory: Forster et al., "On-manifold preintegration for real-time visual-inertial odometry", 2016
http://rpg.ifi.uzh.ch/docs/TRO16_forster.pdf
- Library with preintegrated factors: gtsam.org
- Example system with preintegrated factors: visual-inertial ORB-SLAM
<https://arxiv.org/pdf/1610.05949.pdf>

11. Monocular Camera Localization in 3D LiDAR Maps

Advanced

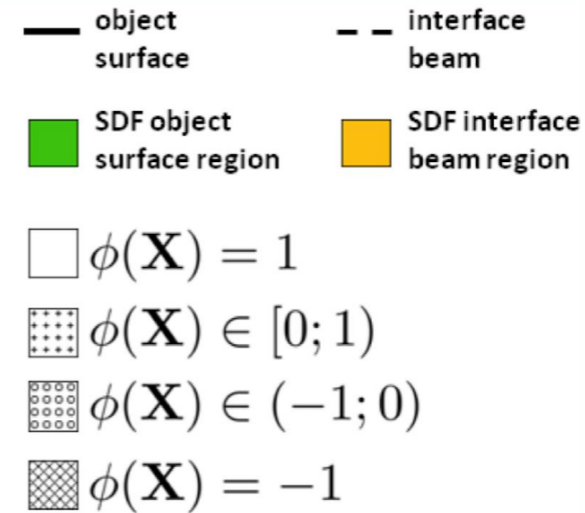
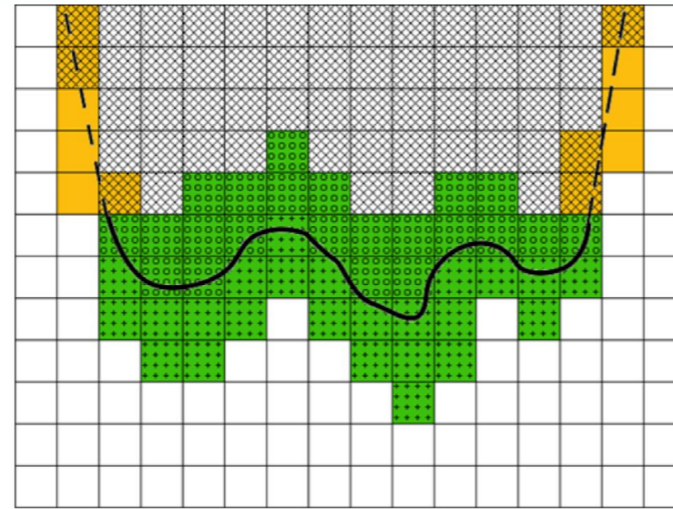
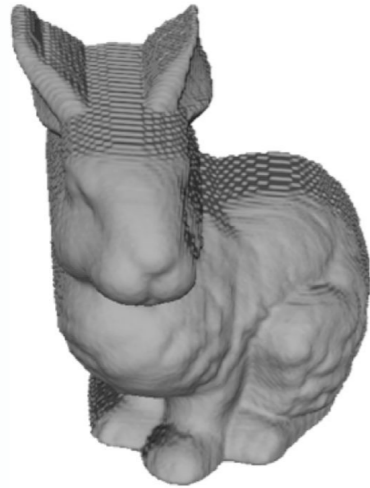


- Localize camera with using pre-built Lidar map
- use ICP(3D-3D pose estimation)
- Voxelized the Lidar map
- avoid boundary effect through mapdistrition

— T. Caselitz, B. Steder, M. Ruhnke and W. Burgard, "Monocular camera localization in 3D LiDAR maps," *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 1926-1931, doi: 10.1109/IROS.2016.7759304.

4. SDF-based Tracking and Reconstruction

Advanced



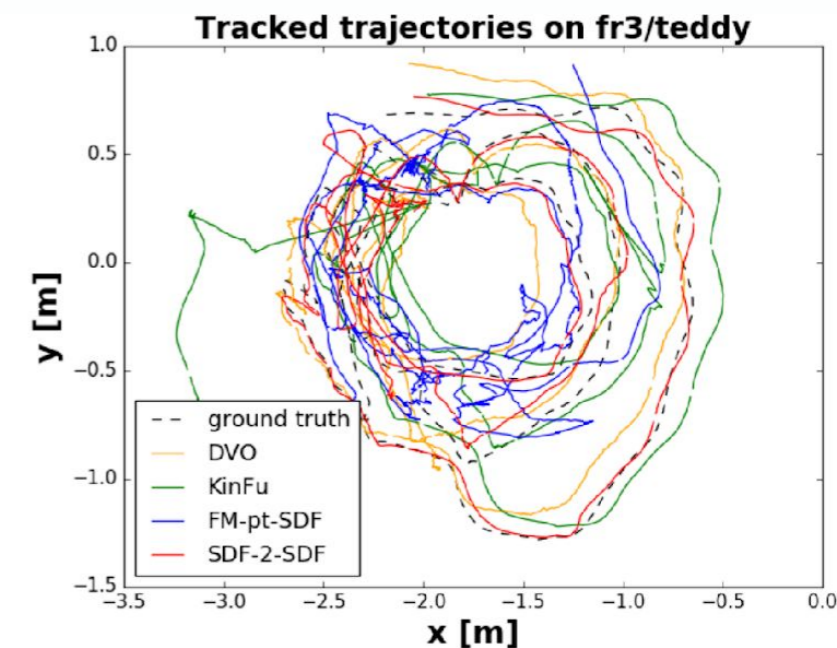
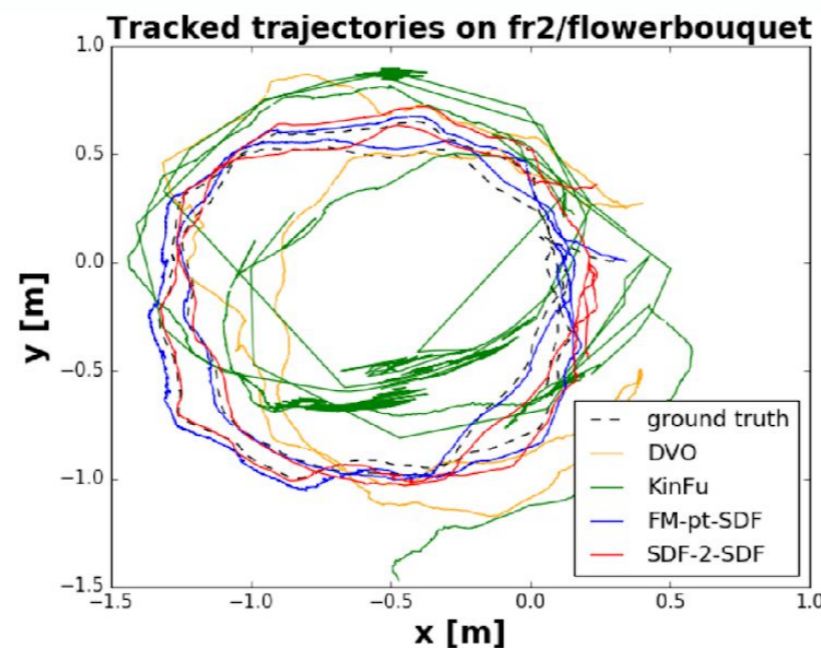
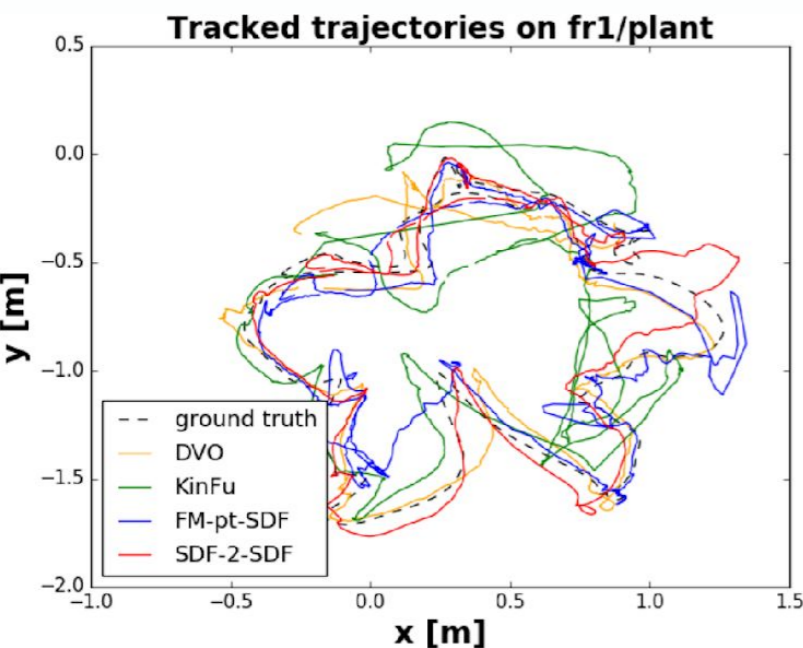
- Signed distance function (SDF) for surface representation
- Implement SDF-based tracking in reconstruction pipeline

• Slavcheva et al., "Sdf-2-sdf: Highly accurate 3d object reconstruction", ECCV 2016

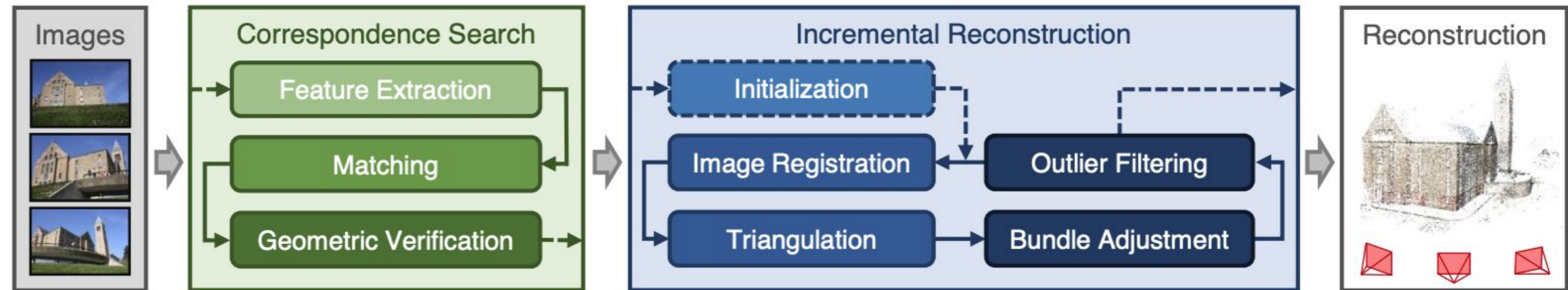
<http://campar.in.tum.de/pub/slavcheva2016eccv/slavcheva2016eccv.pdf>

• Bylow et al., Real-time camera tracking and 3D reconstruction using signed distance functions, RSS 2013

https://vision.cs.tum.edu/_media/spezial/bib/bylow_etal_rss2013.pdf



9. Structure from Motion Revisited



- Improve different stages of SfM application
 - Initialisation
 - Next-best-view selection
 - Triangulation
 - Re-triangulation and Outlier Filtering

- Schonberger, Johannes L., and Jan-Michael Frahm. "Structure-from-motion revisited." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2016.

https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Schonberger_Structure-From-Motion_Revisited_CVPR_2016_paper.pdf

10. SfM for Large Image Collections

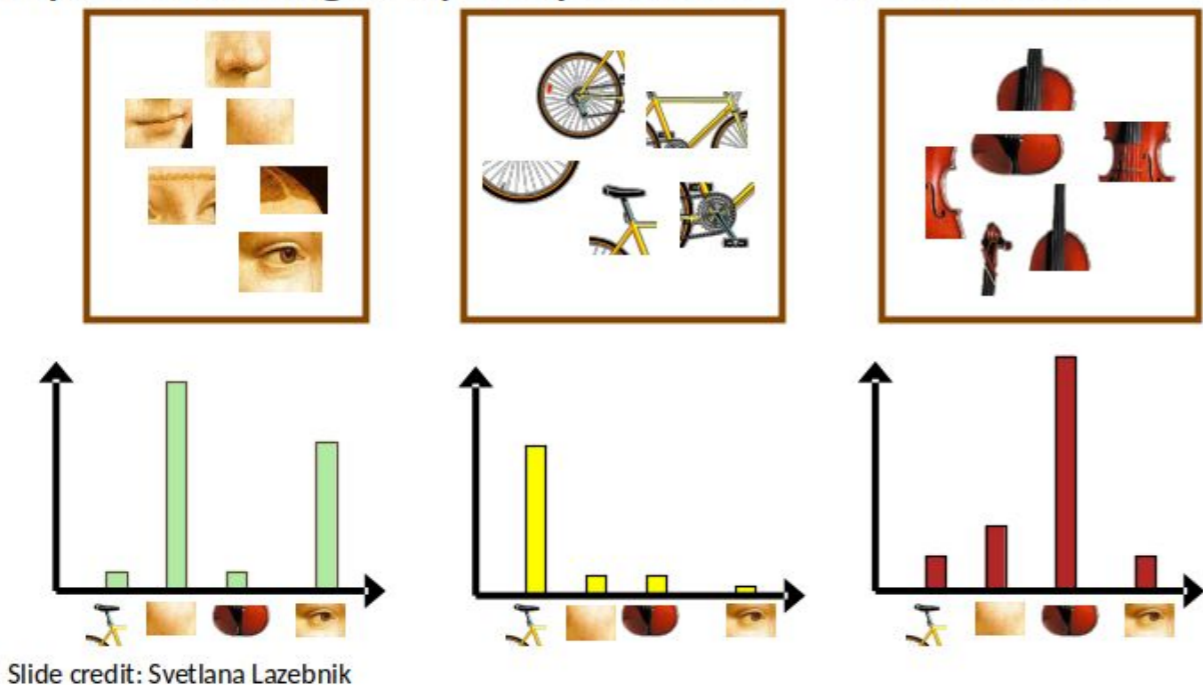
Advanced



- Make SfM work for thousands of images
 - Full Euroc sequences or other datasets
 - Select subset of images
 - Subsample
 - Discard too close or similar images
 - Use BoW for candidate selection
 - Implement direct index for more efficient feature matching
 - More efficient geometric verification
 - Schönberger J.L., Price T., Sattler T., Frahm J.M., Pollefeys M. A Vote-and-Verify Strategy for Fast Spatial Verification in Image Retrieval. ACCV 2016. (<https://demuc.de/papers/schoenberger2016vote.pdf>)

old - Bag of Words for Place Recognition

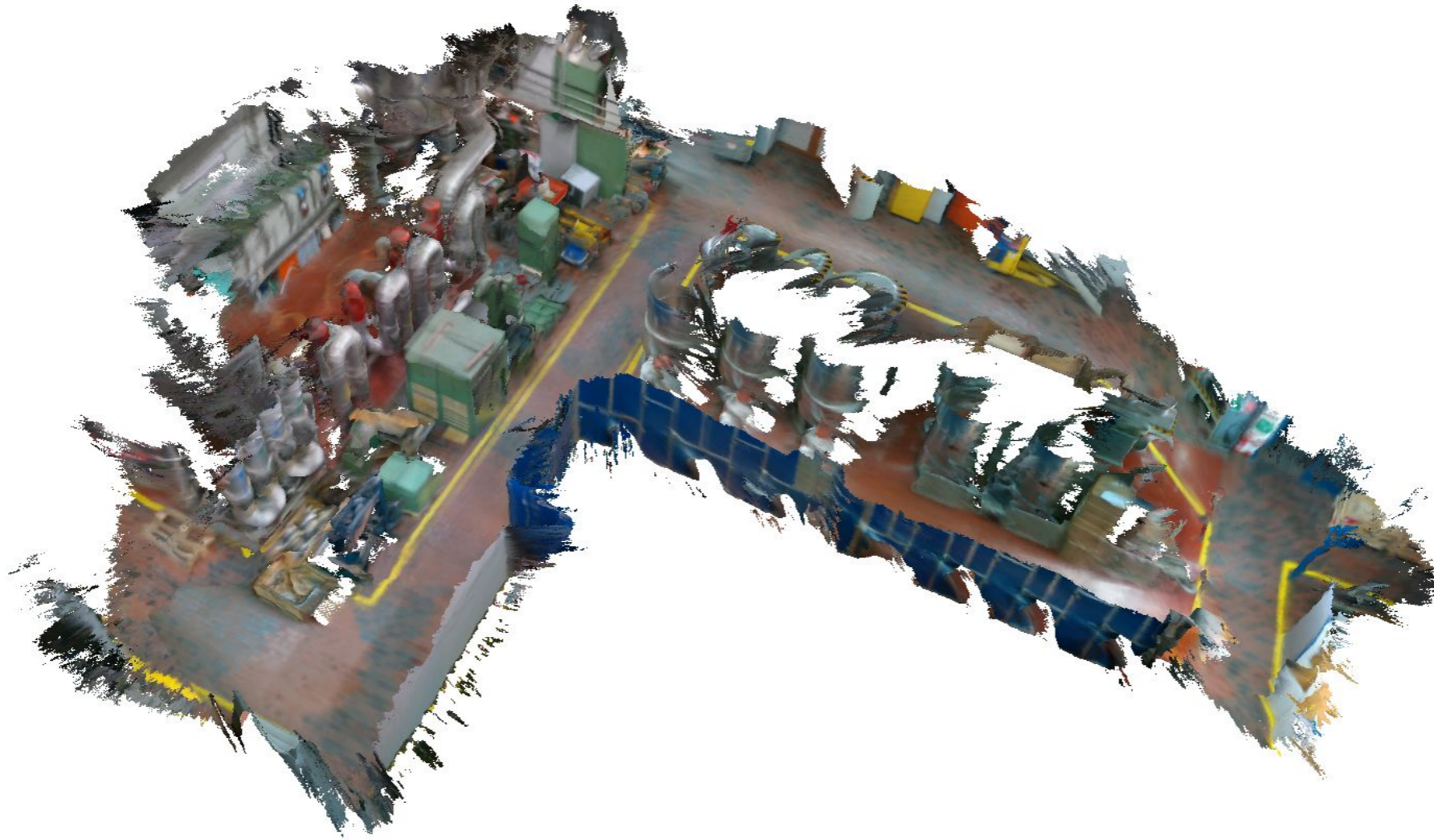
1. Extract local features
2. Learn “visual vocabulary”
3. Quantize local features using visual vocabulary
4. Represent images by frequencies of “visual words”



- Applications:
 - SfM: Speed up pairwise matching
 - SLAM: re-localization, loop-closure

- DBoW2/3
 - Paper: <http://doriangalvez.com/papers/GalvezTRO12.pdf>
 - Code: <https://github.com/rmsalinas/DBow3>
- HBST: A Hamming Distance embedding Binary Search Tree
 - Paper: <https://arxiv.org/abs/1802.09261>
 - Code: https://gitlab.com/srrg-software/srrg_hbst

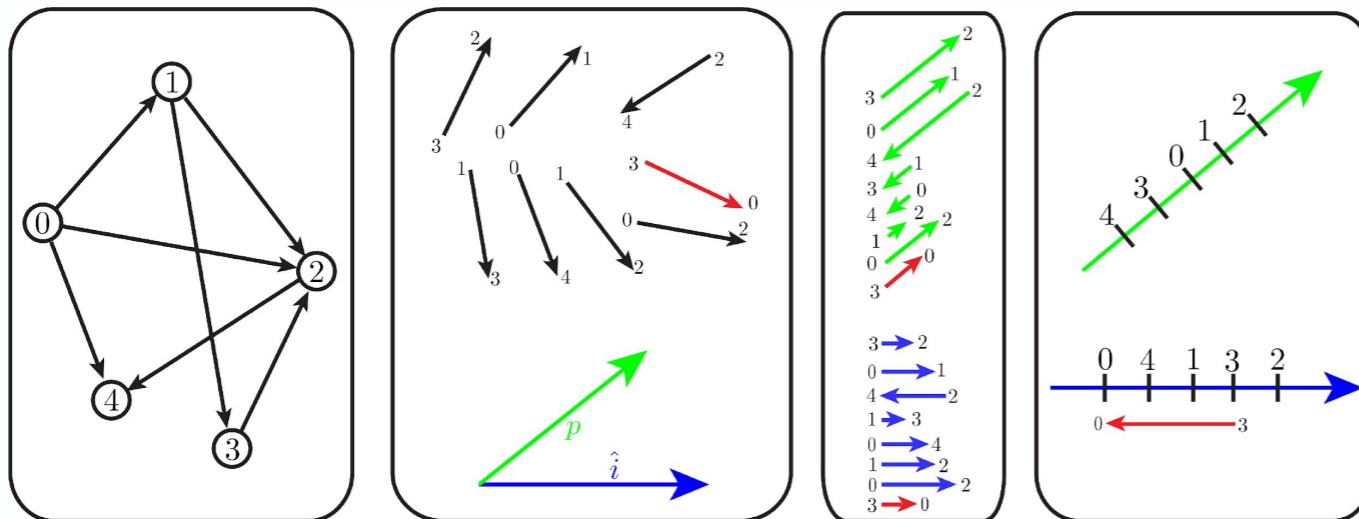
old - Dense 3D reconstruction for SfM



- PSMNet
 - <https://arxiv.org/abs/1803.08669>
 - <https://github.com/JiaRenChang/PSMNet>
- Depth images Fusion
 - Voxblox: <https://github.com/ethz-asl/voxblox>

6. Global SfM with Motion Averaging

- Goal: Implement global SfM pipeline using Motion Averaging (as opposed to the incremental pipeline from sheet 4)
- Approach:
 - Estimate relative rotation between pairs of cameras
 - Solve for global camera orientations
 - Given the global orientations, estimate global translations
 - Triangulate structure



- Chatterjee, Avishek, and Venu Madhav Govindu. "Efficient and robust large-scale rotation averaging." Proceedings of the IEEE International Conference on Computer Vision. 2013.
https://www.cv-foundation.org/openaccess/content_iccv_2013/papers/Chatterjee_Efficient_and_Robust_2013_ICCV_paper.pdf
- Wilson, Kyle, and Noah Snavely. "Robust global translations with 1dsfm." European Conference on Computer Vision. Springer, Cham, 2014.
https://research.cs.cornell.edu/1dsfm/docs/1DSfM_ECCV14.pdf
- Zhu, Siyu, et al. "Very large-scale global sfm by distributed motion averaging." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.
http://openaccess.thecvf.com/content_cvpr_2018/papers/Zhu_Very_Large-Scale_Global_CVPR_2018_paper.pdf