UiO : Department of Technology Systems
University of Oslo

Lecture 10.3
ORB-SLAM

Trym Vegard Haavardsholm

Content and illustrations from
R. Mur-Artal, J. M. M. Montiel, and J. D. Tardos,
“ORB-SLAM: A Versatile and Accurate Monocular SLAM System,”
IEEE Trans. Robot., vol. 31, no. 5, pp. 1147–1163, Oct. 2015.

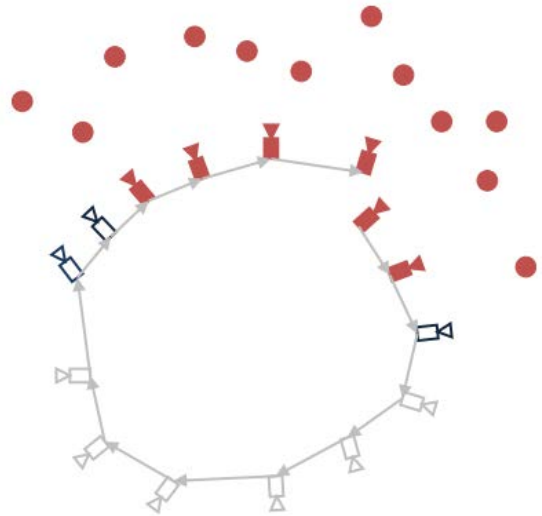
Code: https://github.com/raulmur/ORB_SLAM2

Main contributions

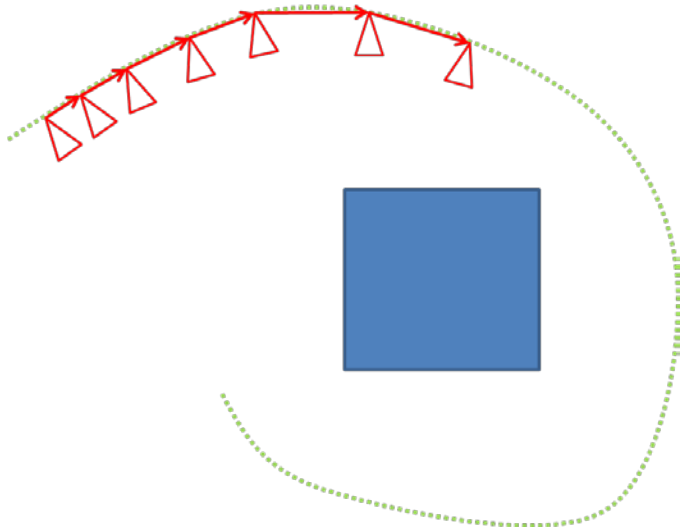
- Use of the same features for all tasks:
  tracking, mapping, relocalization, and loop closing

- Real-time operation in large environments.
  Thanks to the use of a co-visibility graph,
  tracking and mapping are focused in a local co-visible area,
  independent of global map size

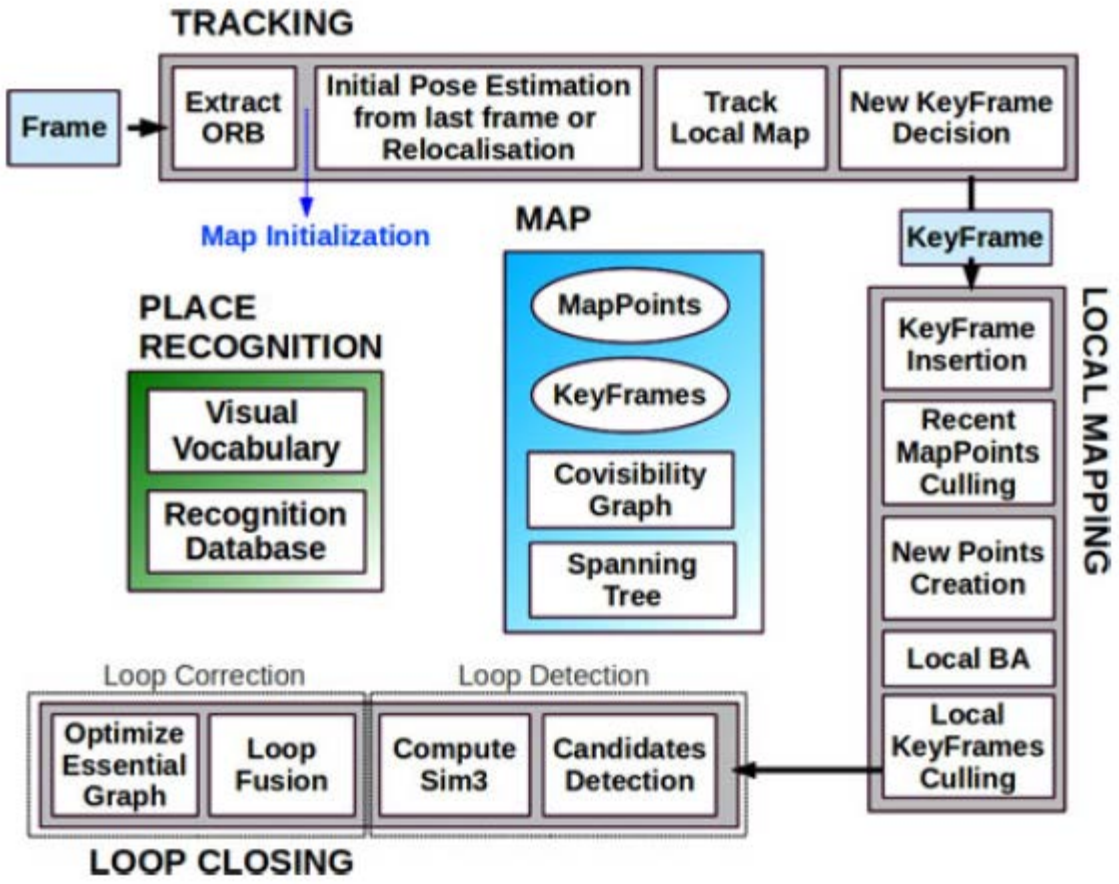
- Real-time loop closing based on the optimization of a pose graph that we call the Essential Graph



Strasdat, H., Davison, A. J., Montiel, J. M. M., & Konolige, K. (2011).
Double window optimisation for constant time visual SLAM.
Proceedings of the IEEE International Conference on Computer Vision, 2352–2359



System overview



Short-term tracking

- FAST corners in grid cells at different scale levels with ORB descriptors

- Initial pose estimation:
  - Tracking OK: Guided search with constant velocity model
  - Tracking ~OK: Wider search around positions in previous frame
  - Tracking lost: Global relocalization (long-term tracking)

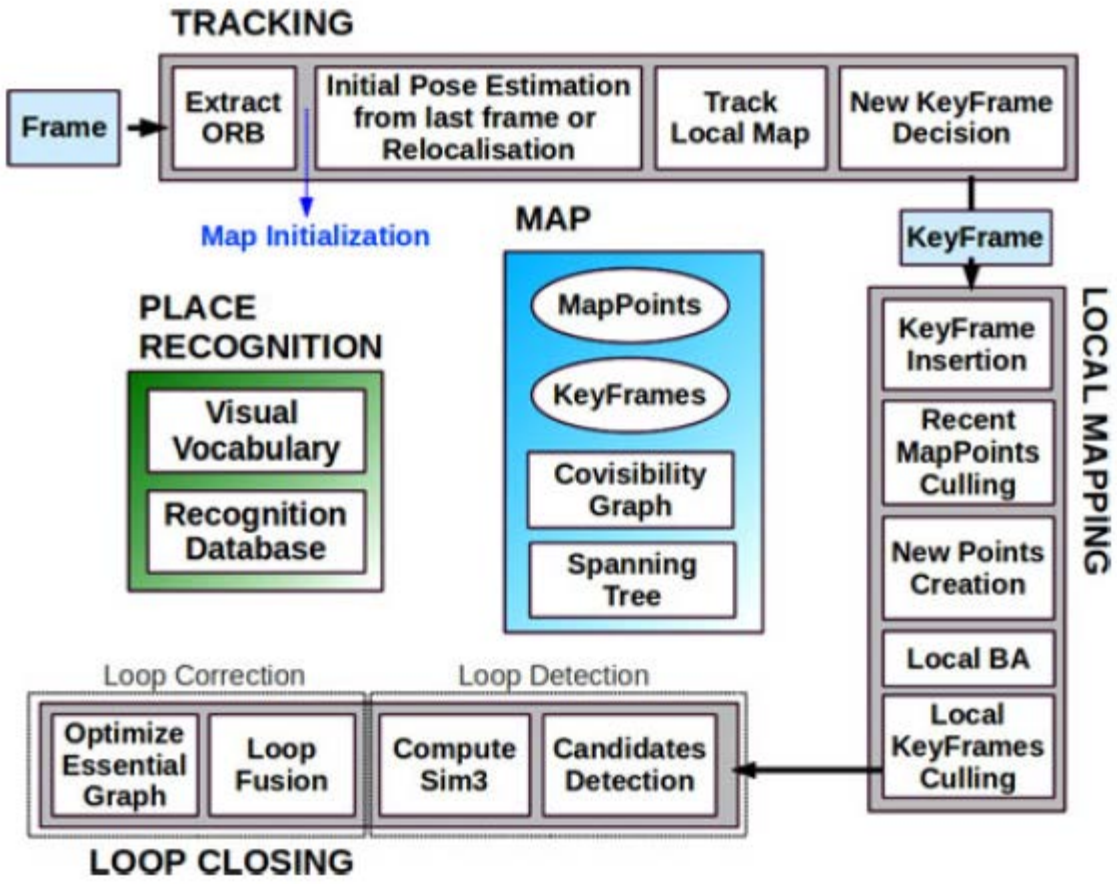
- Track local map
  - Project local map and search for more correspondences
  - Motion-only bundle adjustment

New keyframe decision

Insert keyframes often to make tracking more robust to rotations

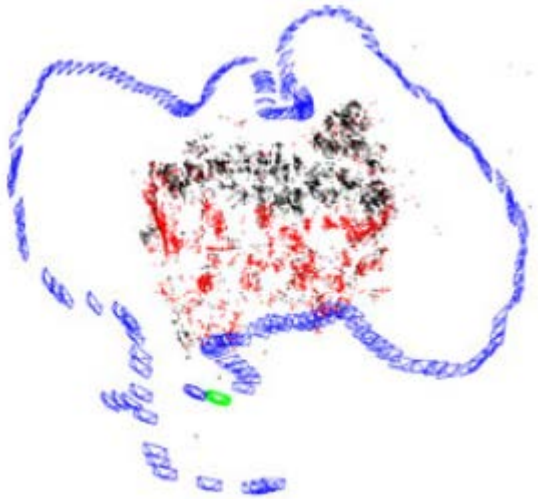
- More than 20 frames since last global relocalization
- Local mapping is idle,
  or more than 20 frames since last keyframe insertion
- Current frame tracks at least 50 points
- Current frame tracks less than 90% than that of the reference keyframe

Mapping


TRACKING
Frame
Extract ORB
Initial Pose Estimation from last frame or Relocalisation
Track Local Map
New KeyFrame Decision
Map Initialization
MAP
KeyFrame
PLACE RECOGNITION
Visual Vocabulary
Recognition Database
MapPoints
KeyFrames
Covisibility Graph
Spanning Tree
KeyFrame Insertion
Recent MapPoints Culling
New Points Creation
Local BA
Local KeyFrames Culling
LOCAL MAPPING
Loop Correction
Loop Detection
Optimize Essential Graph
Loop Fusion
Compute Sim3
Candidates Detection
LOOP CLOSING

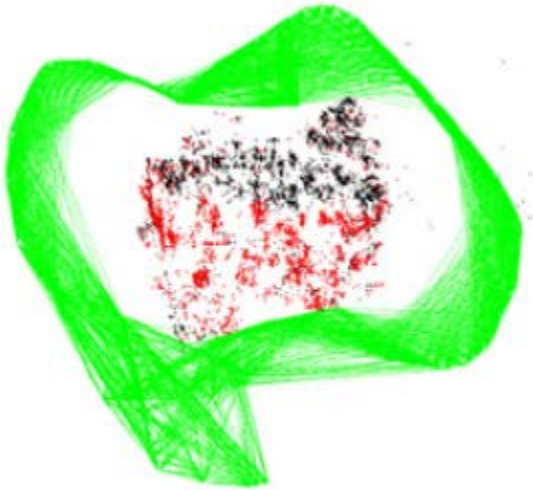
Map

- Keyframes (blue)
- Current frame (green)
- Map points (black)
- Active map points (red)



Map

- Co-visibility graph
  - Nodes: All keyframes
  - Edges: Number of common map points (at least 15)

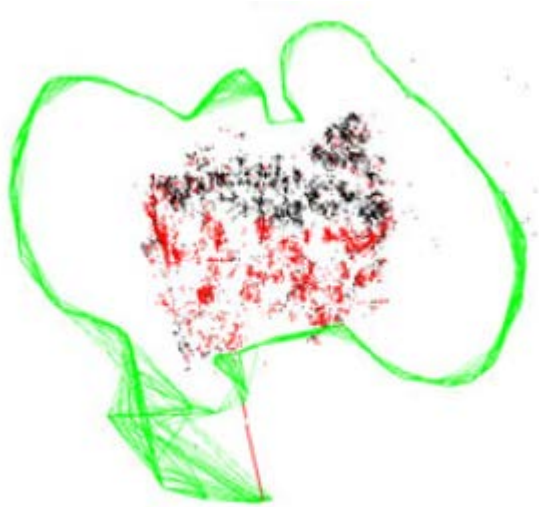


Map

- Spanning tree
  - Connected subgraph of the co-visibility graph with minimal number of strong edges

Map

- Essential graph
  - Spanning tree
  - Subset of edges from the co-visibility graph with high co-visibility (at least 100)
  - Loop closure edges



Mapping

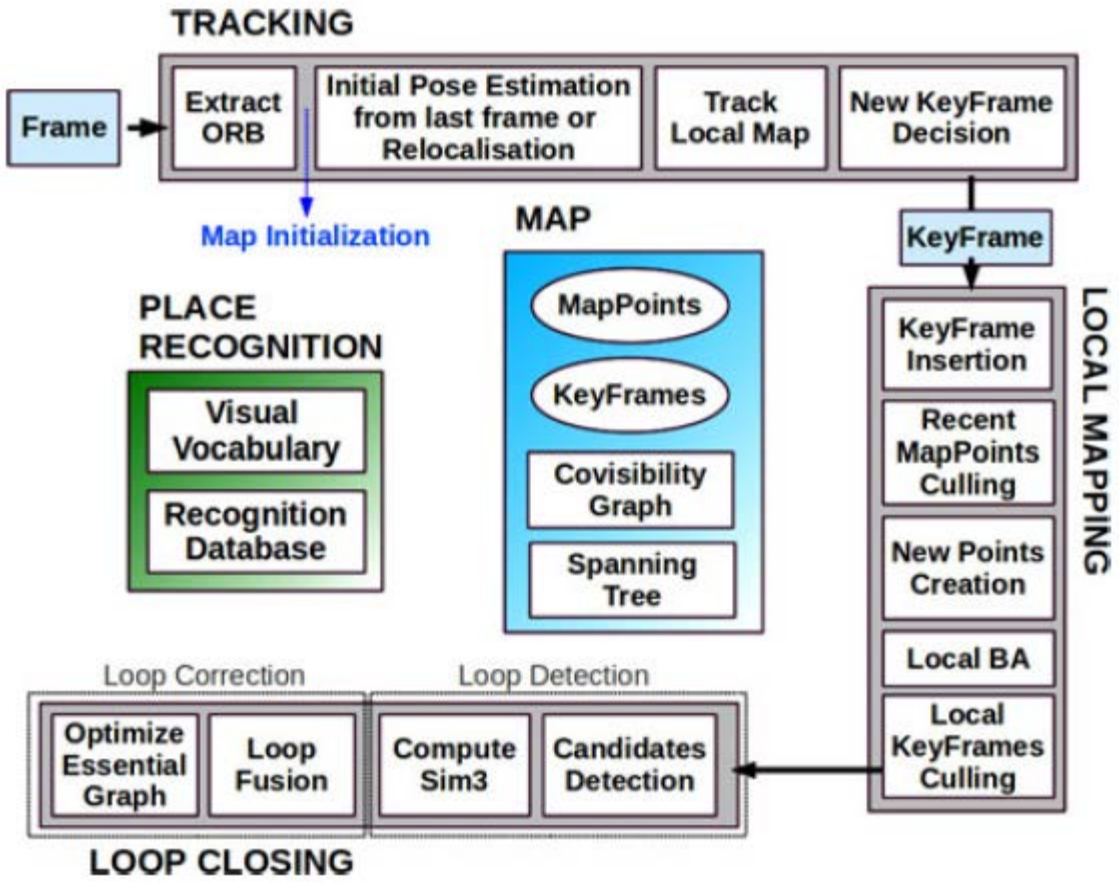
- Keyframe insertion
  - Add to co-visibility graph and spanning tree
- Recent map point culling
  - Remove bad points during the first three keyframes after creation
- New map point creation
  - Triangulate ORB from connected keyframes

Mapping

- Local bundle adjustment
  - Optimize the current keyframe,
    all connected keyframes and all points seen
  - All other keyframes remain fixed

- Local keyframe culling
  - Detect and delete redundant keyframes

Long-term tracking



Long-term tracking

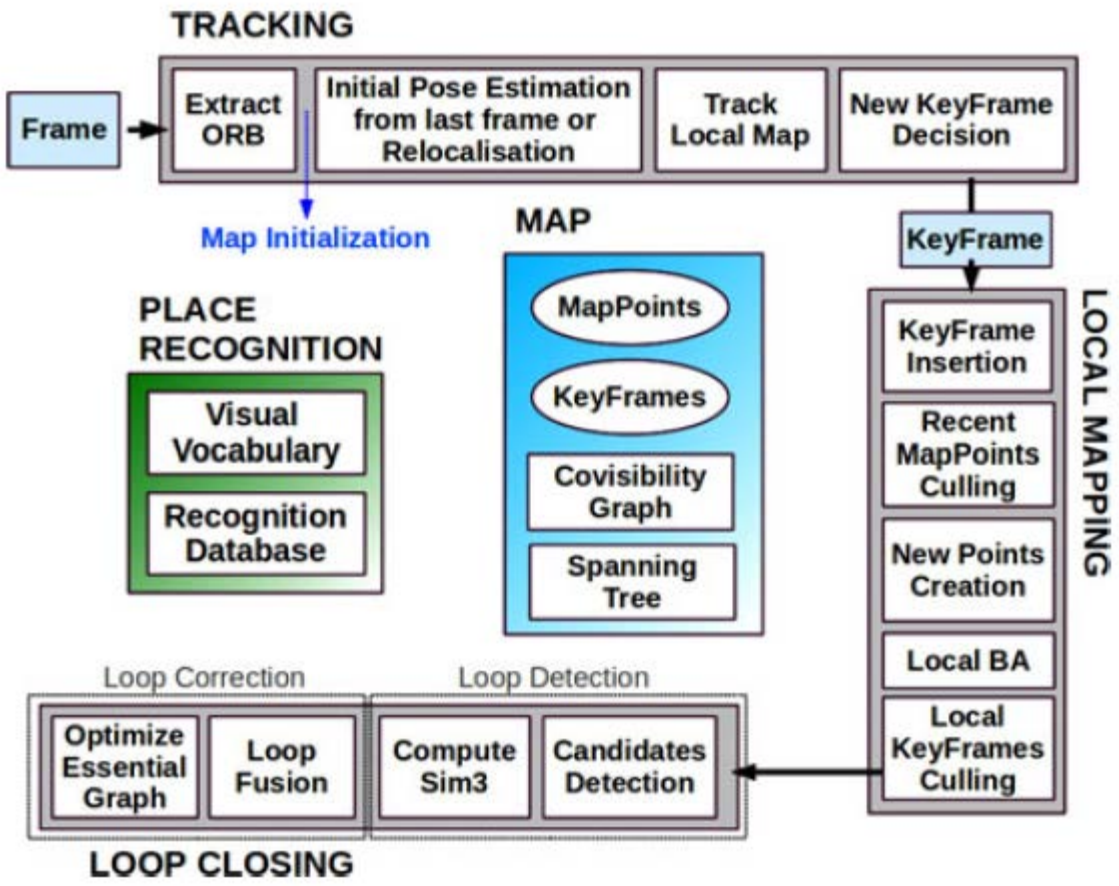
- Long-term tracking is performed on the last keyframe in its own thread

- Loop closure candidates:
  - Compute BoW threshold based on the lowest score for the neighboring keyframes
  - Query recognition database (DBoW2) for keyframes with score higher than the threshold
  - Keep those that are not directly connected, and where we have at least three connected candidates

Long-term tracking

- Scale-drift aware loop closure alignment
  - Compute an initial similarity transform between the current keyframe and the loop keyframe from 3D-to-3D correspondences
  - Search for more correspondences
  - Optimize again
  - Geometric validation: Accept loop if enough inliers

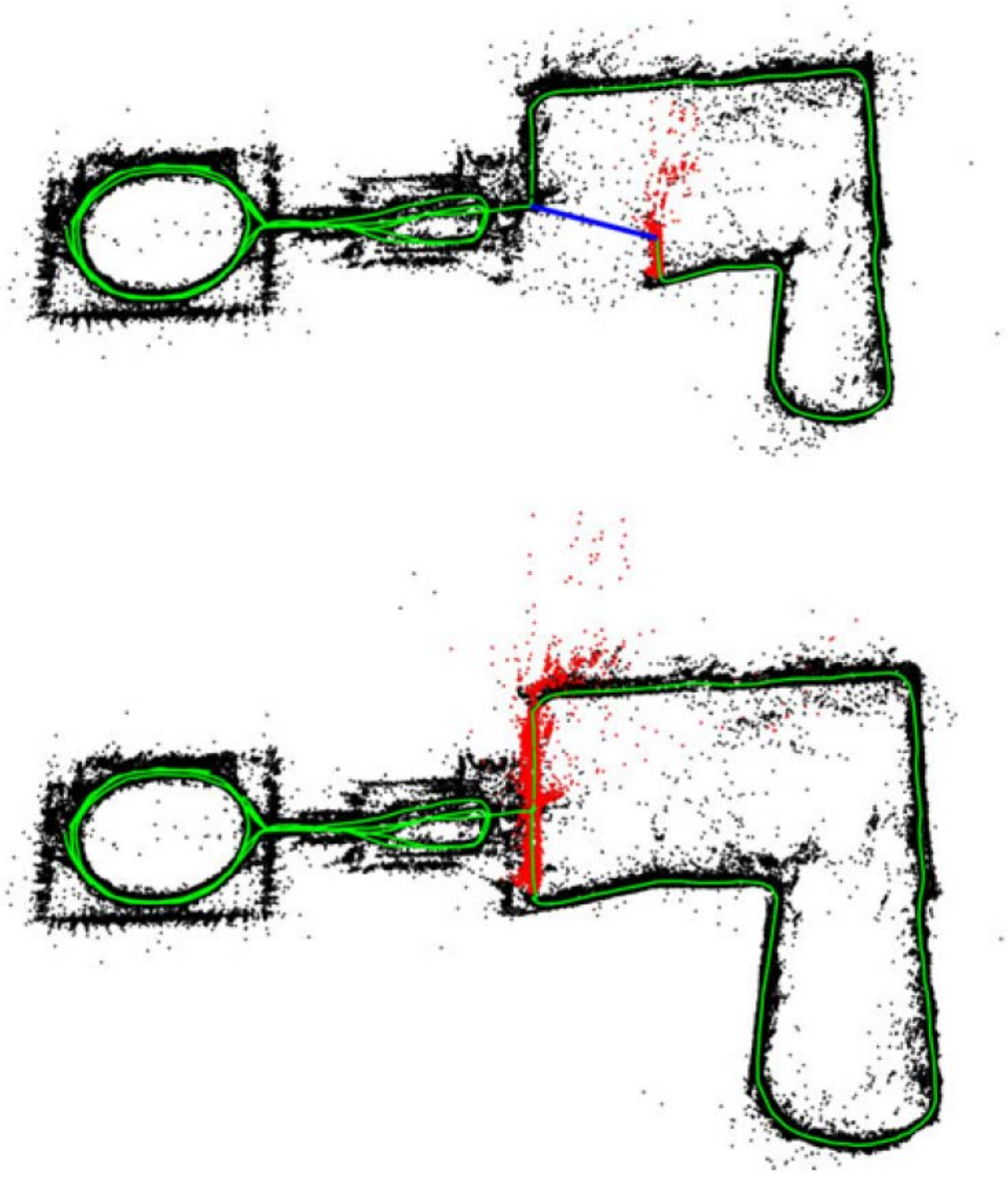
Loop correction



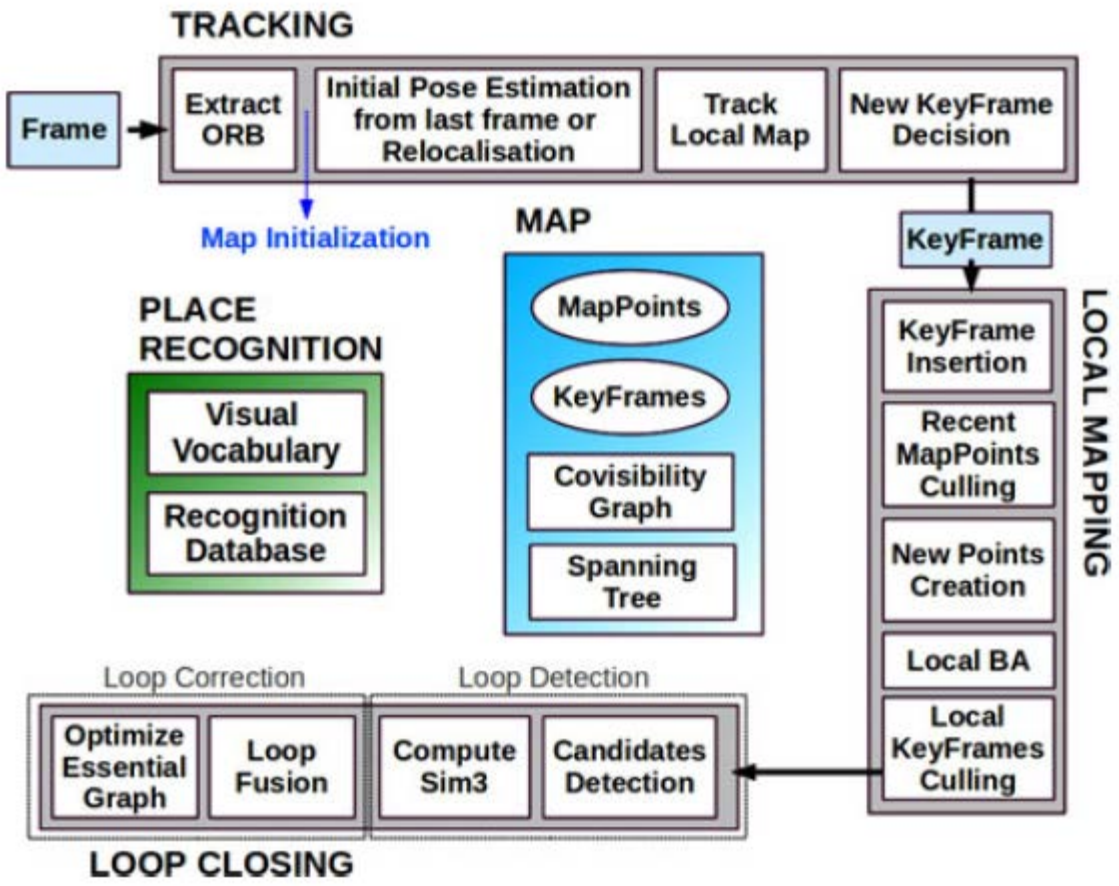
Loop correction

- Loop fusion
  - Fuse map points
  - Insert new edges in the co-visibility graph

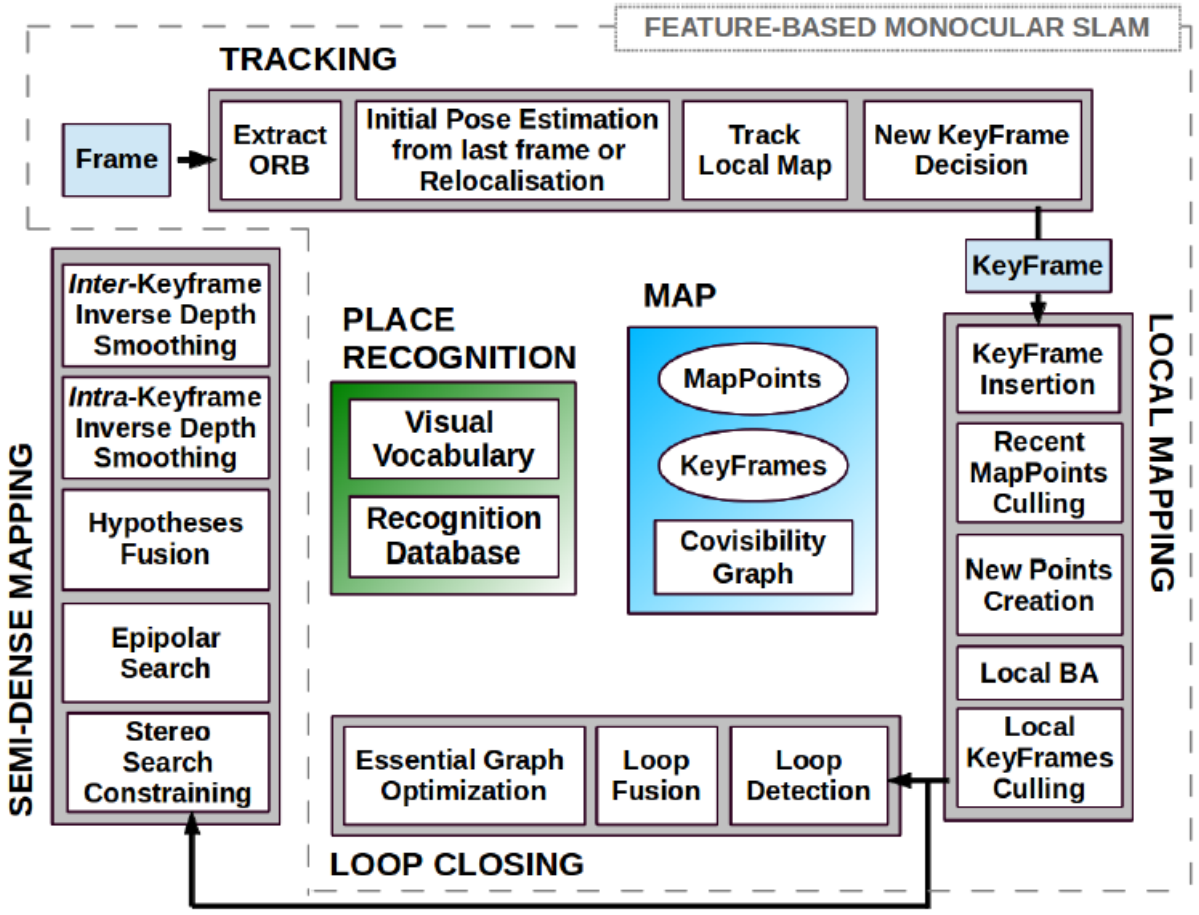
- Essential graph optimization
  - Distribute the loop closing error along a pose graph over $\mathfrak{sim}(3)$
  - Transform each map point according to the correction of one of the keyframes that observes it



Summary

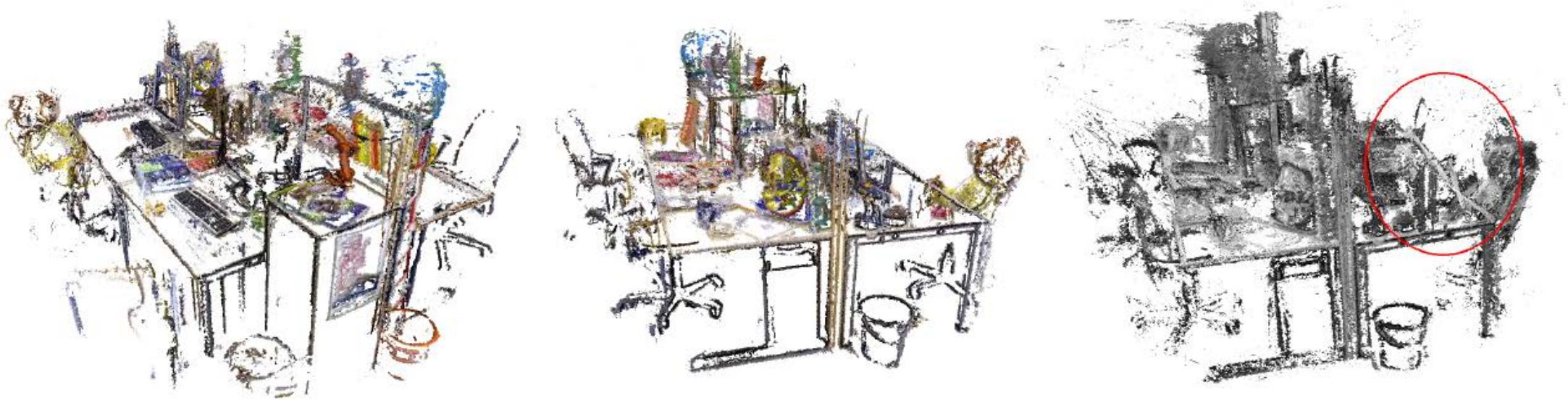

TRACKING
Frame
Extract ORB
Initial Pose Estimation from last frame or Relocalisation
Track Local Map
New KeyFrame Decision
Map Initialization
MAP
MapPoints
KeyFrames
Covisibility Graph
Spanning Tree
KeyFrame
PLACE RECOGNITION
Visual Vocabulary
Recognition Database
LOCAL MAPPING
KeyFrame Insertion
Recent MapPoints Culling
New Points Creation
Local BA
Local KeyFrames Culling
Loop Correction
Loop Detection
Optimize Essential Graph
Loop Fusion
Compute Sim3
Candidates Detection
LOOP CLOSING

Augmentations: Semi-dense mapping



R. Mur-Artal and J. Tardos, “Probabilistic Semi-Dense Mapping from Highly Accurate Feature-Based Monocular SLAM,” in Robotics: Science and Systems XI, 2015.

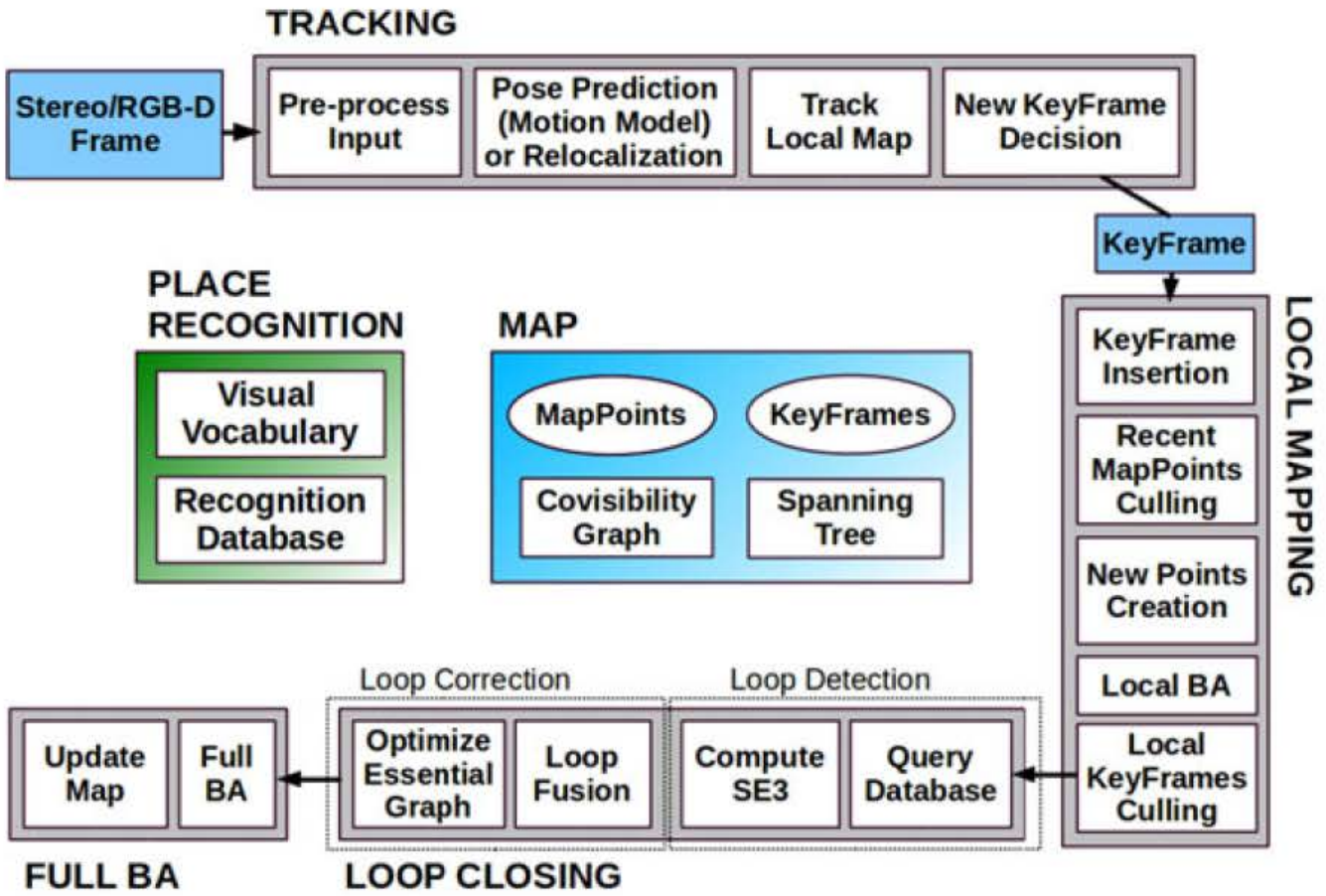
Augmentations: Semi-dense mapping

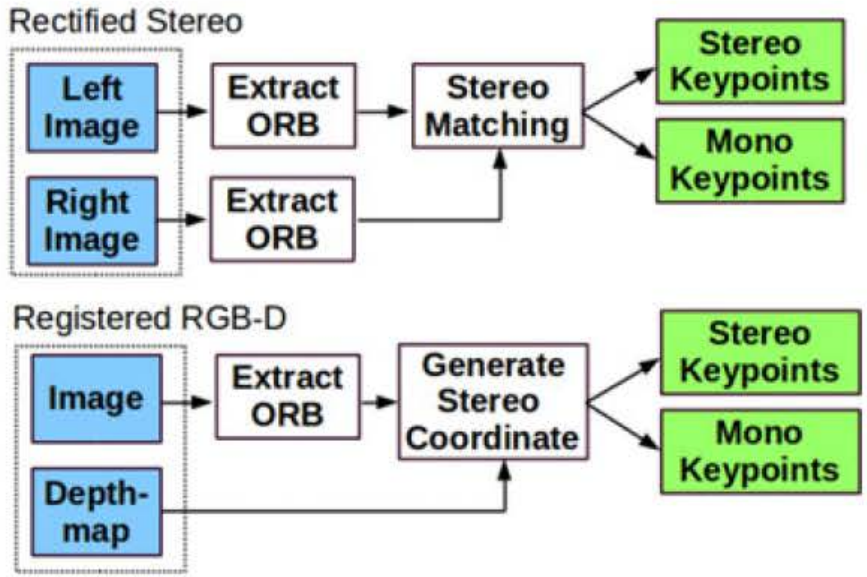


(c) Sequence: *fr3_long_office_household*. Left and Middle: Our system. Rigth: LSD-SLAM

R. Mur-Artal and J. Tardos, “Probabilistic Semi-Dense Mapping from Highly Accurate Feature-Based Monocular SLAM,” in Robotics: Science and Systems XI, 2015.

Augmentations: ORB-SLAM 2





R. Mur-Artal and J. D. Tardos, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," IEEE Trans. Robot., pp. 1–8, 2017.