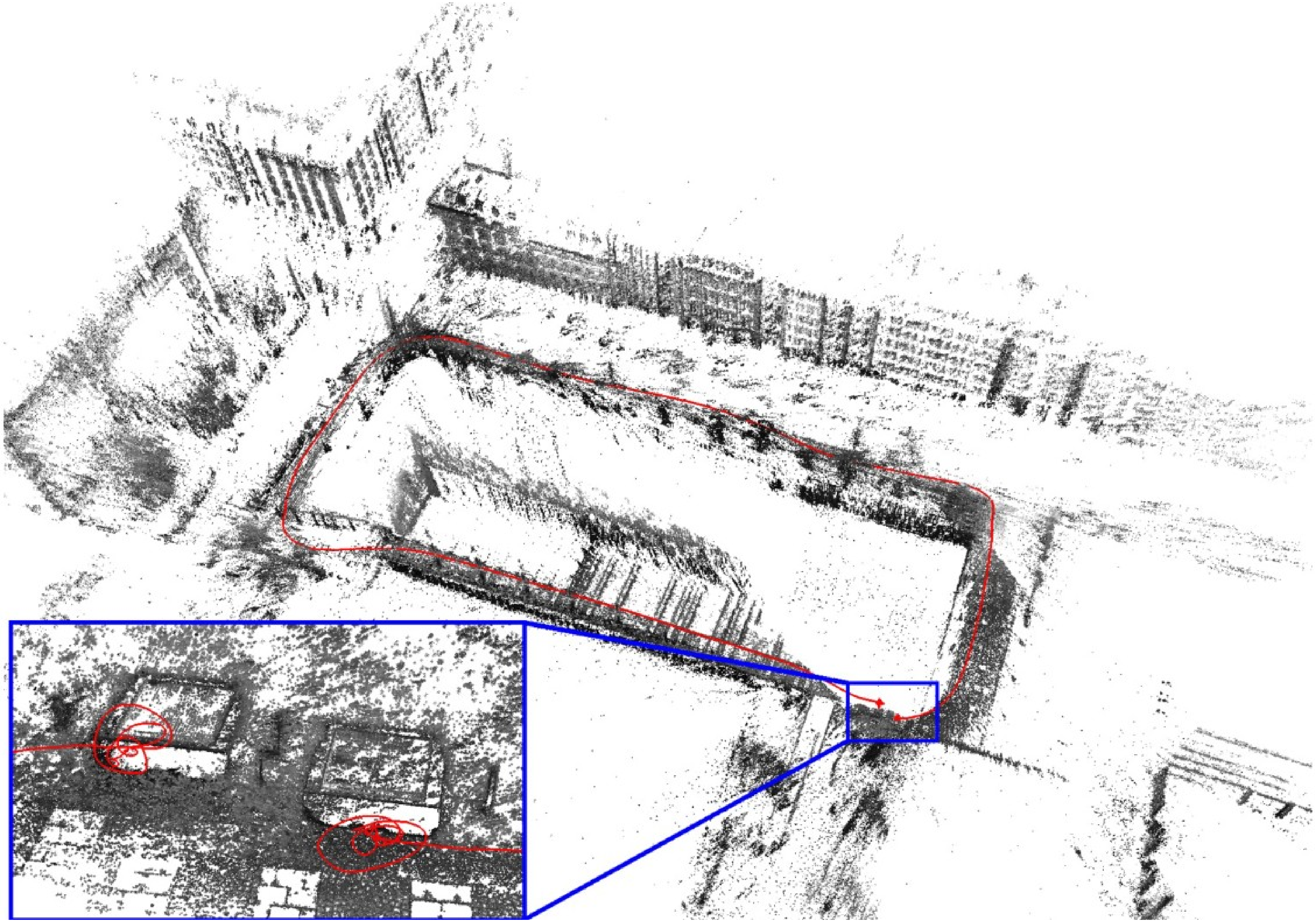
Simultaneous Localization and Mapping

Feature based approach



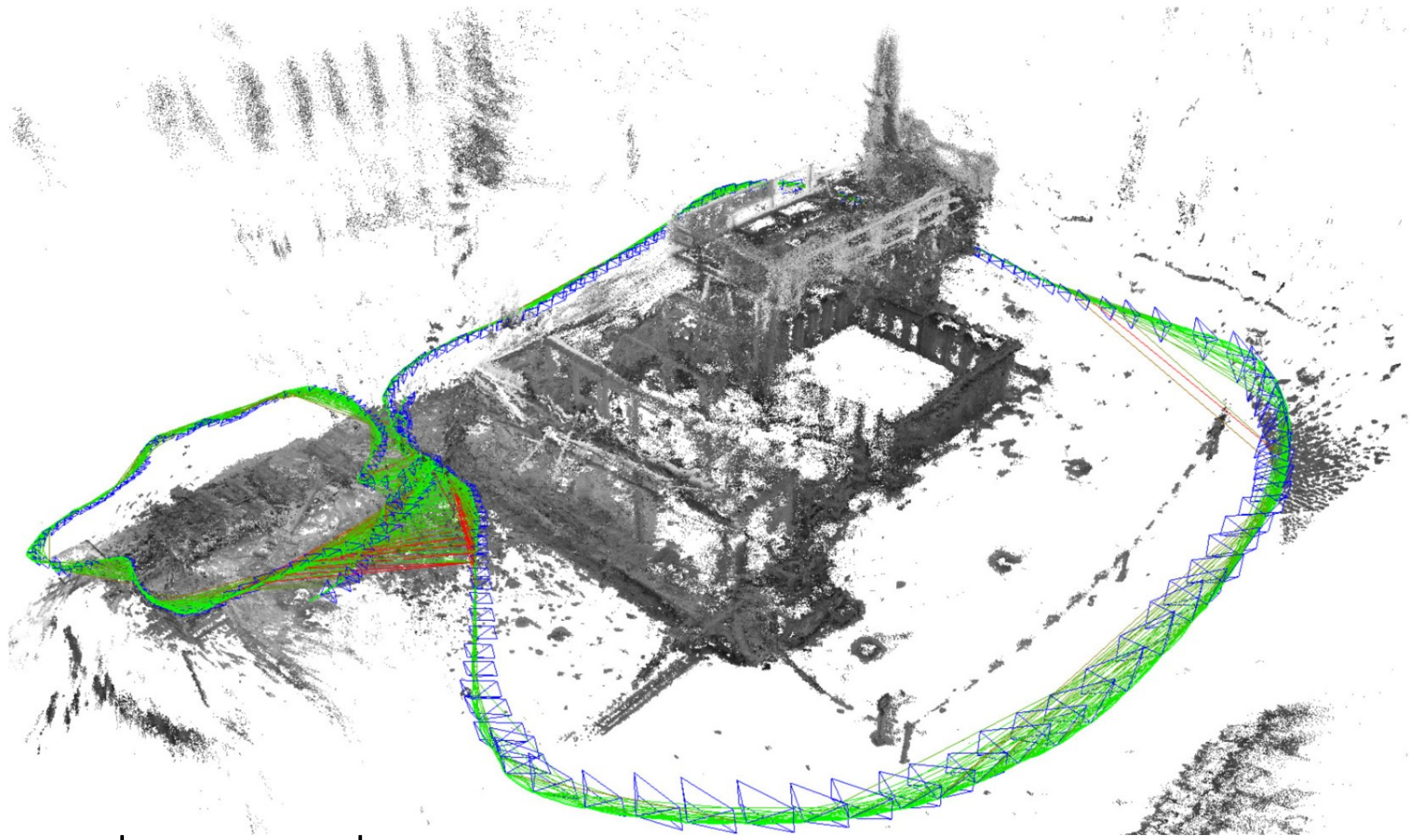
SLAM vs Odometry

- Past: Odometry had no loop closure
- Now: closing gap
- Odometry focuses on localization
- SLAM focuses on both mapping and localization

Measurements, Sensors

- Camera (RGB, RGB-D, TOF, infra, wide-narrow, stereo…)
- Inertial Measurement Unit - IMU
  - accelerometer, gyroscope, magnetometer
- GPS
- Lidar
- Ultrasonic sensor
- Augmented environment (MoCap, AprilTag…)
- Microphone, Rotary sensor, WiFi, LiFi, Bluetooth…

Maps



- Metric vs Topological
- Implicit representations
  - Occupancy grids
  - Depth fields
  - Light fields, radiance fields
- Explicit representations
  - Point clouds
  - Keyframes
  - Meshes
- Graphs
  - Pose graph, factor graph, covisibility graph, scene graph…
- Grids
  - Voxel grid, Multi resolution, Hierarchical, Octree, k-d tree

Basic formulation

- Probabilistic model
  - **Y** noisy measurements
  - **X** unknown model parameters (map, trajectory)
- Maximum Likelihood approach
  - X_{opt} = argmax P(**Y**|**X**)
  - Only the (approx.) best solution
- Maximum A Posteriori approach
  - P(**X**|**Y**) = P(**Y**|**X**) * P(**X**) * c
  - Complete distribution over the parameters
  - Prior often unfeasible

EKF vs Keyframe based SLAM

- EKF: Extended Kalman Filter
  - Strict probabilistic approach
  - Hard to detect/incorporate loop closures
  - Marginalization is difficult
  - Better suited to factor graphs
- Keyframe based
  - Sparser approach
  - Easier long-term association for loop closure and bundle adjustment
  - More robust

V/VI-SLAM

- Visual vs Visual+Inertial Odometry
- Most common
- IMU measurements
  - More information, higher accuracy
  - More complexity
- Differences
  - Pose graph optimization
  - Initialization, calibration
  - Local/global Bundle adjustment

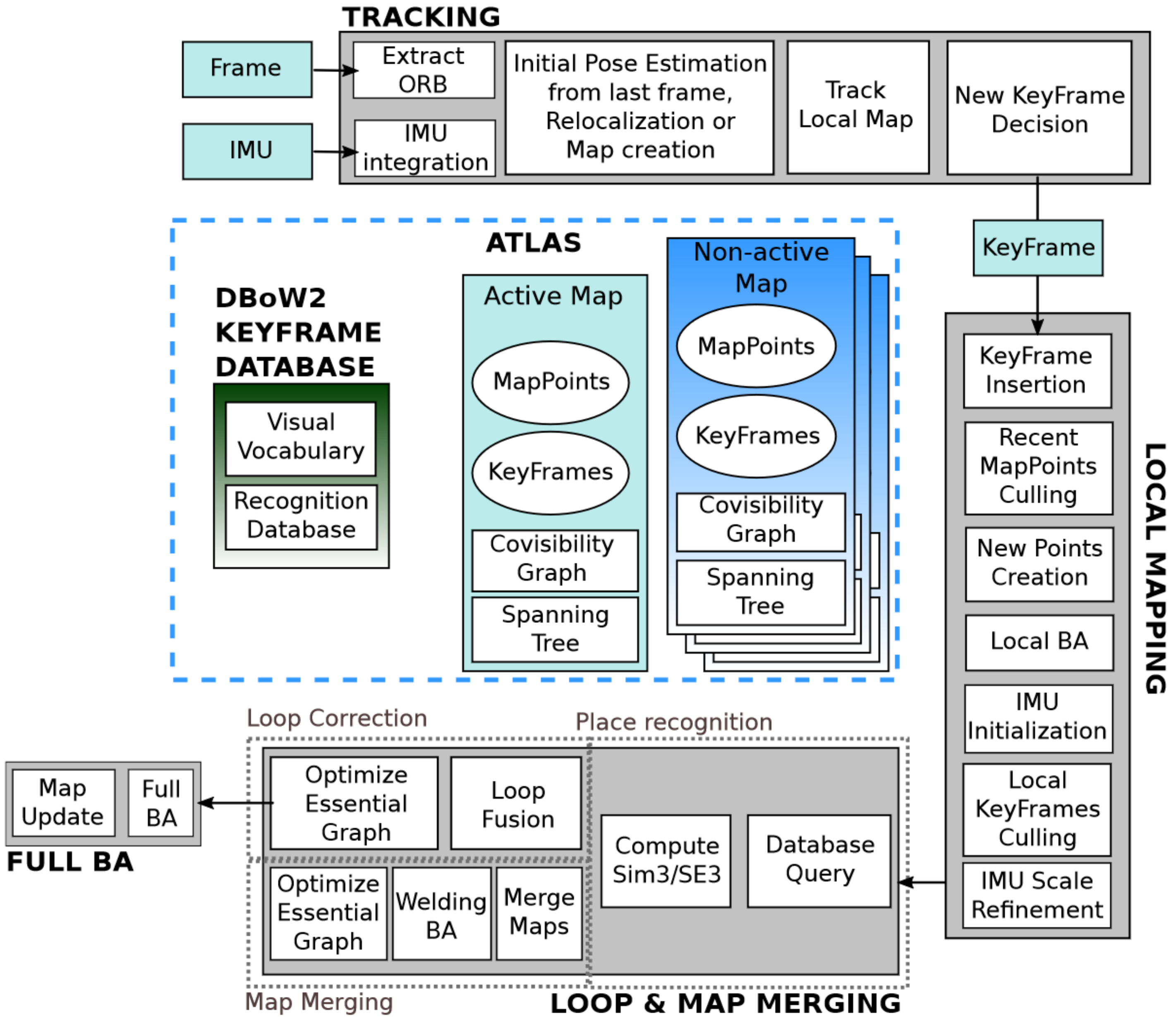
V/VI-SLAM Approaches

- Direct vs Indirect
  - Indirect: raw measurements preprocessed -> **Y**
    (e.g. features, optical flow, line detection)
    typically geometric error
  - Direct: light (radiant energy or radiance) as **Y**
    typically photometric error (geometric for depth measurements)
- Dense vs Sparse
  - Dense: all pixels are used during the estimation
    keeps geometrical prior: notion of neighborhood, leads to dense Hessians
  - Sparse: "special" pixels are selected (corners, line segments)
    keypoint positions conditionally independent given the camera parameters
  - (Semi-dense: not all pixels, but larger patches)

Sparse + Indirect

- Most common
- Map
  - Keyframes, keyframe descriptors
  - Feature points (2D and 3D) with descriptors
  - Pose graph
- Localization
  - Feature point extraction
  - Feature/frame descriptor generation
  - Image retrieval
  - Feature matching
  - Ransac+PnP

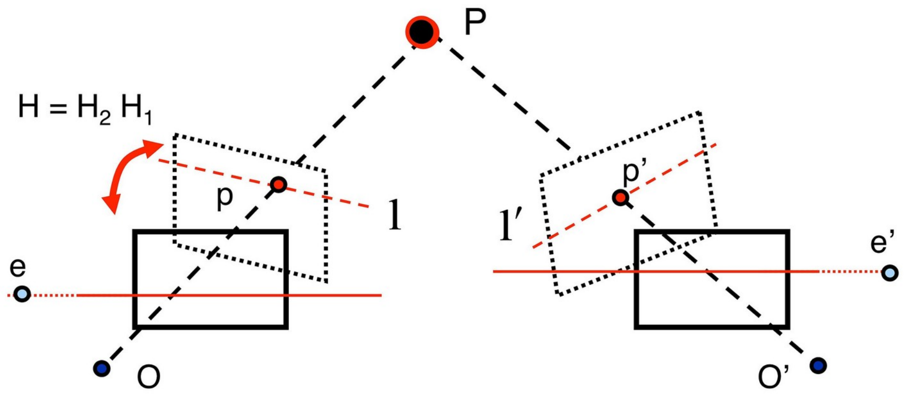
Basic steps



TRACKING
Frame
IMU
Extract ORB
IMU integration
Initial Pose Estimation from last frame, Relocalization or Map creation
Track Local Map
New KeyFrame Decision
KeyFrame
ATLAS
DBoW2 KEYFRAME DATABASE
Visual Vocabulary
Recognition Database
Active Map
MapPoints
KeyFrames
Covisibility Graph
Spanning Tree
Non-active Map
MapPoints
KeyFrames
Covisibility Graph
Spanning Tree
LOCAL MAPPING
KeyFrame Insertion
Recent MapPoints Culling
New Points Creation
Local BA
IMU Initialization
Local KeyFrames Culling
IMU Scale Refinement
Loop Correction
Place recognition
Optimize Essential Graph
Loop Fusion
Compute Sim3/SE3
Database Query
Optimize Essential Graph
Welding BA
Merge Maps
Map Merging
LOOP & MAP MERGING
Map Update
Full BA
FULL BA

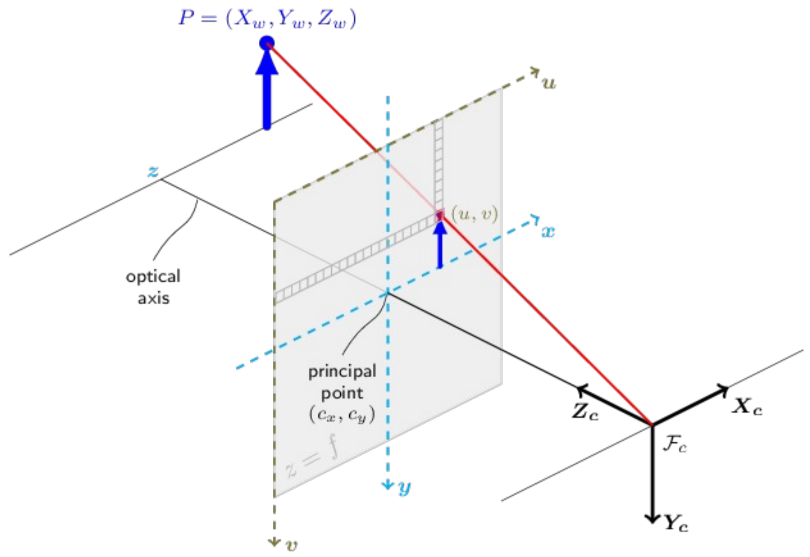
Tracking – Taking new measurements

- Undistort image
- Color, exposure balancing…
- Rectify (stereo): epipolar lines parallel
- Pinhole camera model



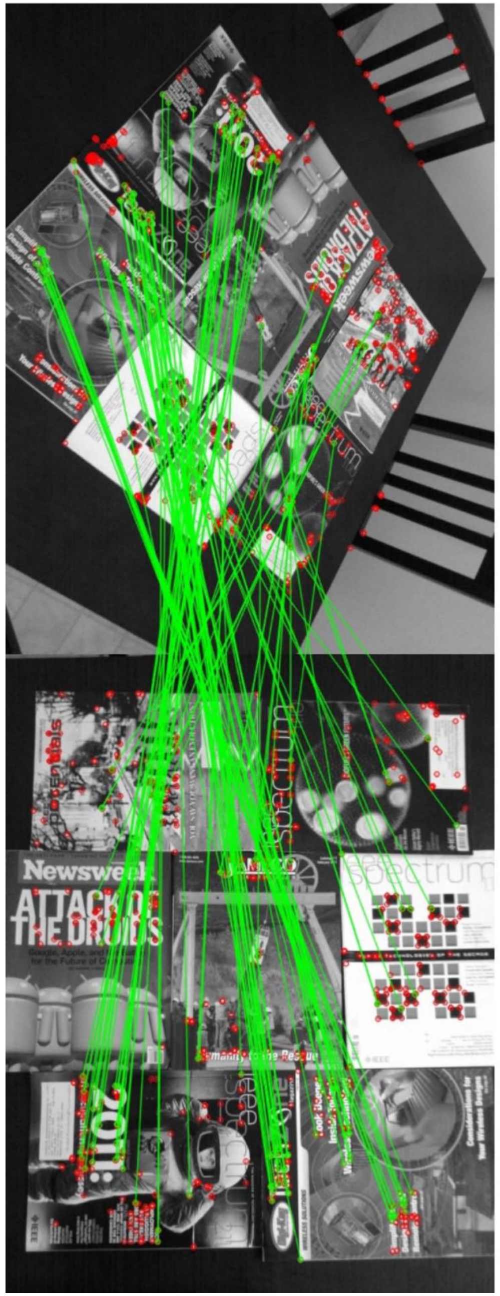
$$s \begin{bmatrix} u \\ v \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x \\ 0 & f_y & c_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_x \\ r_{21} & r_{22} & r_{23} & t_y \\ r_{31} & r_{32} & r_{33} & t_z \end{bmatrix} \begin{bmatrix} X_w \\ Y_w \\ Z_w \\ 1 \end{bmatrix}$$

$$(k_1, k_2, p_1, p_2[, k_3[, k_4, k_5, k_6[, s_1, s_2, s_3, s_4[, \tau_x, \tau_y]]]])$$



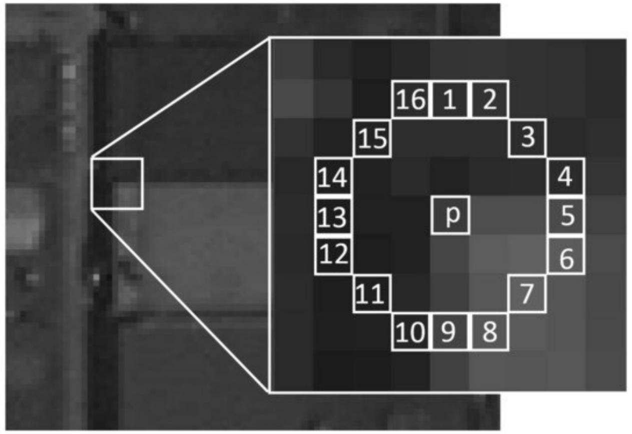
Tracking- Feature extraction

- High gradient points: edges, corners
- Optionally multi scale and oriented
- Shi-Tomasi, Harris corner detection: fast, inaccurate
- FAST: fast, single scale, not oriented
- SIFT: slowest, multi scale, oriented, patented
- SURF: slow, multi scale, oriented (inaccurate), patented
- ORB: fast, multi scale, oriented, free



ORB: Oriented Fast and Rotated BRIEF

- FAST threshold for circular ring around center -> involves edges too
- Harris corner filtering, top N points
- Scale pyramid for multi scale
- Orientation from center of mass
- BRIEF (Binary robust independent elementary feature)
  - Binary intensity tests in the patch
- Rotate tests according to feature orientation
- Use greedy algorithm to find best test pairs



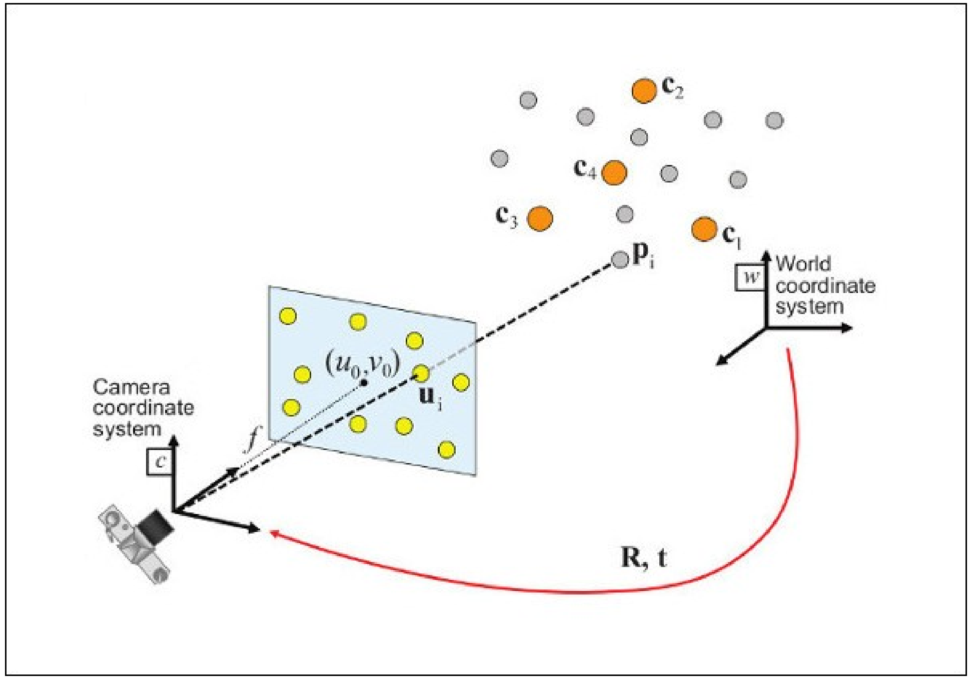
$$m_{pq} = \sum_{x,y} x^p y^q I(x,y)$$

$$C = \left(\frac{m_{10}}{m_{00}}, \frac{m_{01}}{m_{00}} \right)$$

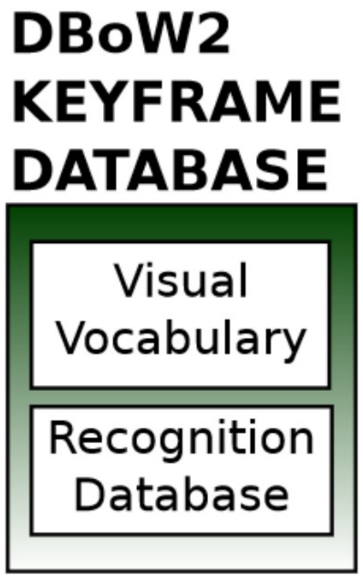
$$\theta = \operatorname{atan2}(m_{01}, m_{10})$$

Tracking- Pose estimation

- Current keyframe with feature points
- Feature matching based on desctiptors (cosine similarity)
- RANSAC + PnP
- Inliers, outliers
- If fails
  - relocalization
- If succeeds
  - Keyframe optimization



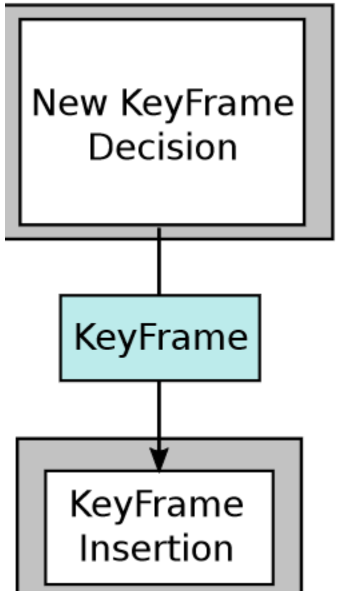
Relocalization



- Input: map, measurements
- Output: **T** pose estimation
- Keyframe based approach:
  - Image retrieval with BoW
  - Pose estimation with feature matching
- DBoW2: Bags of Binary Words for Fast Place Recognition in Image Sequence
  - General visual vocabulary
  - Inverted index
  - Updated recognition database
  - Multiple solutions

Tracking- Keyframe selection

- Based on heuristics
  - Time, distance, failed tracking, unbalanced map, feature density…
- Previous keyframe fixed
- New keyframe added to map



Local Mapping – Keyframe insertion

- Update Pose graph
- Calculate BoW descriptor
- Find covisible keyframes
- Match features points
- Discard duplicate feature points
- Triangulate depth
- Project 3D keypoints

Local Mapping – Local Bundle Adjustment

- New keyframe optimized (camera pose Sim(3) or SE(3), 3D features)
- Based on correspondences
- Moving window of keyframes
- Projection (j-th image to i-th)
- Energy term
- Loss function
- Reprojection error (geometric)
- Covariance as weight
  associated with feature scale
- Usually first-order approximations
  (Levenberg–Marquardt alg.)

$$\pi_i(\mathbf{T}_{iw}, \mathbf{X}_{w,j}) = \begin{bmatrix} f_{i,u} \dfrac{x_{i,j}}{z_{i,j}} + c_{i,u} \\ f_{i,v} \dfrac{y_{i,j}}{z_{i,j}} + c_{i,v} \end{bmatrix}$$

$$\begin{bmatrix} x_{i,j} & y_{i,j} & z_{i,j} \end{bmatrix}^T = \mathbf{R}_{iw} \mathbf{X}_{w,j} + \mathbf{t}_{iw}$$

$$\mathbf{e}_{i,j} = \mathbf{x}_{i,j} - \pi_i(\mathbf{T}_{iw}, \mathbf{X}_{w,j})$$

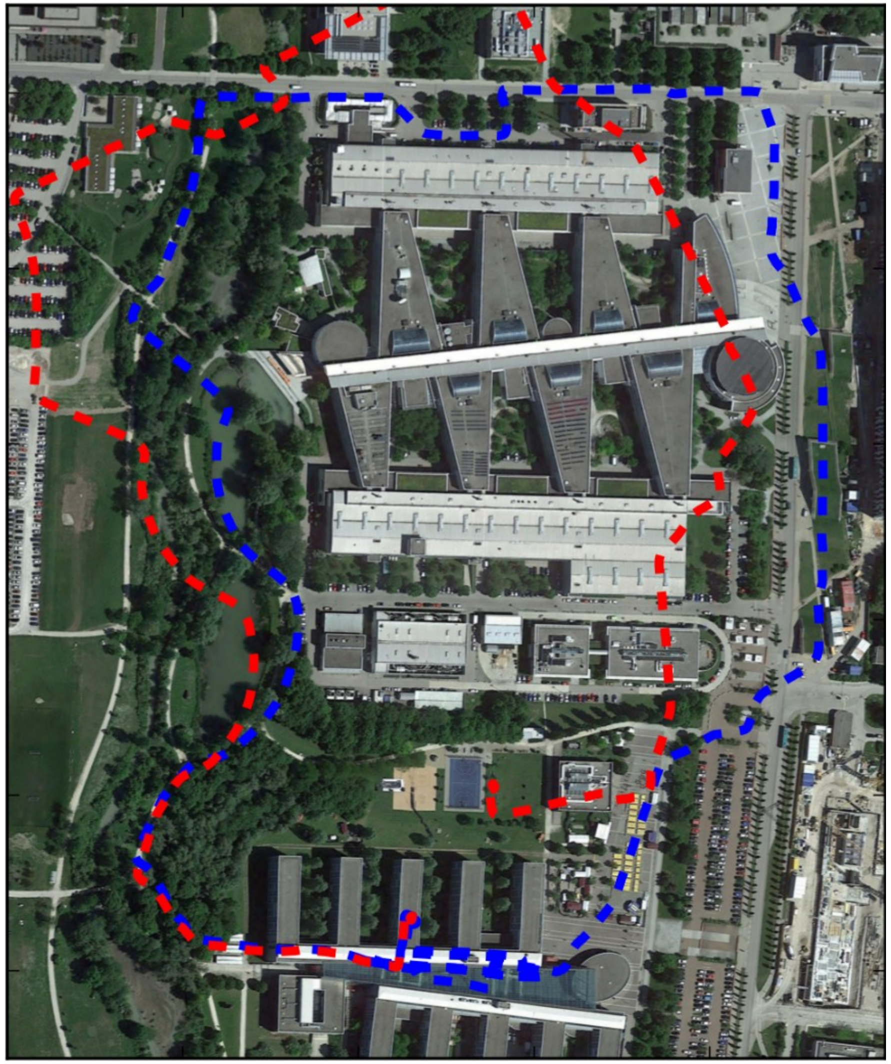
$$C = \sum_{i,j} \rho_h(\mathbf{e}_{i,j}^T \mathbf{\Omega}_{i,j}^{-1} \mathbf{e}_{i,j})$$

Local Mapping – Keyframe culling

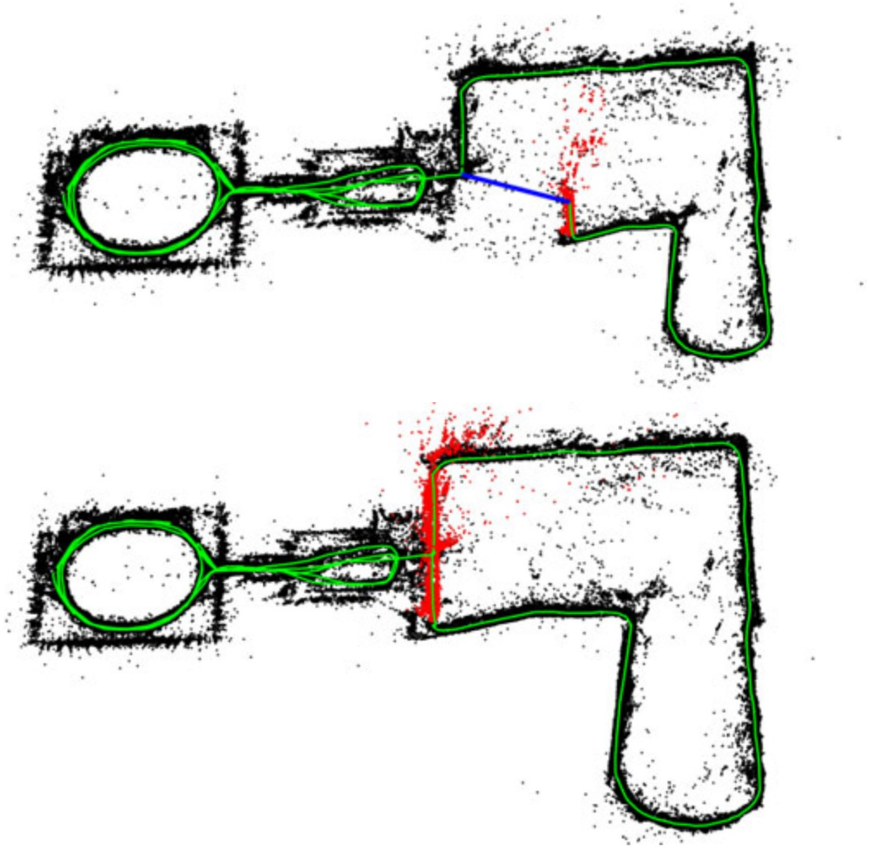
- Discard duplicated keyframes, feature points
- Balance map density
- Keep current environment densely mapped to help tracking
- Sparsify later

Drift

- Odometry contains inaccuracies
- Even with IMU with perfect calibration
  - dead reckoning
- GPS helps but not accurate enough
- EKF solutions suffer from drift as well
- Limitation: maximum mid-term data association
- Global consistency not enforced



Loop Closing – Loop detection



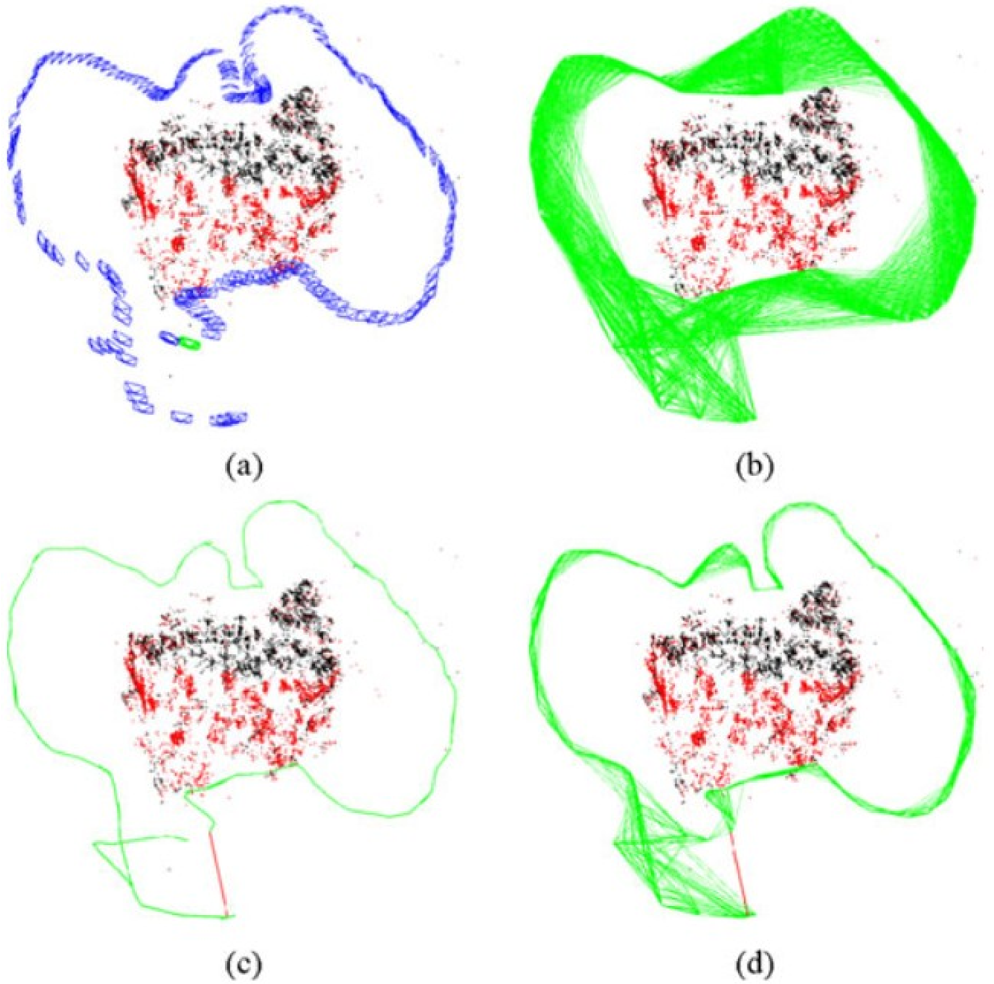
- BoW
- Cosine similarity
- Multiple candidates, local window
- Inlier based verification
- Estimate relative **T** (Sim(3) or SE(3)) transformation
  - E.g. Ransac + Horn algorithm (3D to 3D)
- Refine relative pose with feature matching-based optimization
- Merge keyframes
- Update factor graph or pose graph optimization
- Full BA

Loop Closing – Pose graph optimization

- Optimize only the keyframe poses
- Optionally expressed by a factor graph (preferred if uncertainty is modelled)

$$\mathbf{e}_{i,j} = \log_{\mathrm{Sim}(3)}(\mathbf{S}_{ij}\, \mathbf{S}_{jw}\, \mathbf{S}_{iw}^{-1})$$

$$C = \sum_{i,j} (\mathbf{e}_{i,j}^T \mathbf{\Lambda}_{i,j} \mathbf{e}_{i,j})$$



Full Bundle adjustment

- Formulation similar to Local BA
- Computationally demanding
- All keyframe poses and keypoint positions are optimized

$$\{\mathbf{X}^i, \mathbf{R}_l, \mathbf{t}_l | i \in \mathcal{P}_L, l \in \mathcal{K}_L\} = \underset{\mathbf{X}^i, \mathbf{R}_l, \mathbf{t}_l}{\operatorname{argmin}} \sum_{k \in \mathcal{K}_L \cup \mathcal{K}_F} \sum_{j \in \mathcal{X}_k} \rho(E_{kj})$$

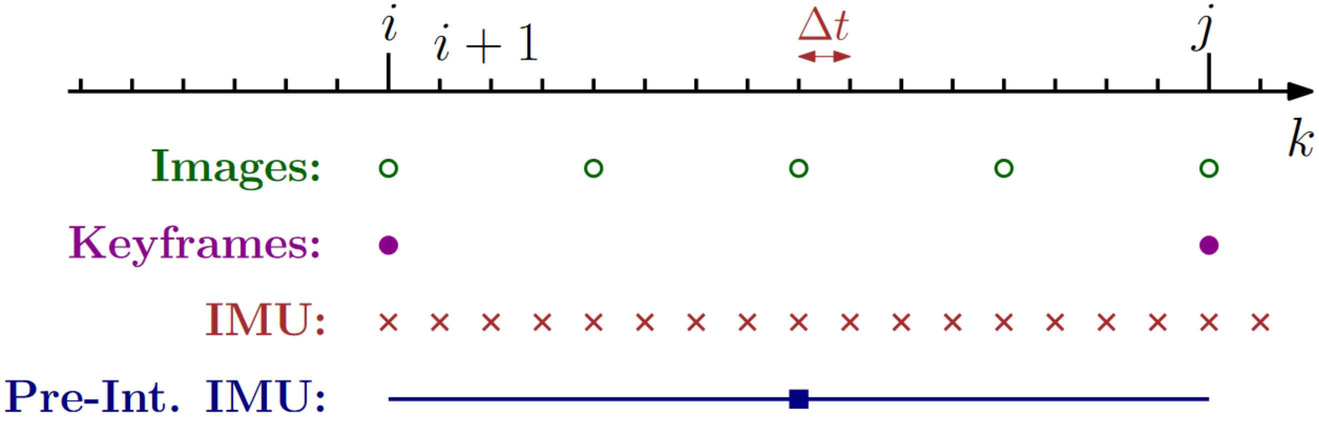
$$E_{kj} = \left\| \mathbf{x}^j_{(\cdot)} - \pi_{(\cdot)}\left(\mathbf{R}_k \mathbf{X}^j + \mathbf{t}_k\right) \right\|^2_{\Sigma}$$

Mono vs Stereo

- Depth uncertainty
- Scale drift
- SE(3) – Sim(3)
- Mono
    - scale calibration
    - scale optimization during pose graph optimization
- Stereo
    - image rectification
    - **Y** contains inverse depth
    - Stereo keypoint: (u_L, v_L, u_R)
    - RGB-D can replace
    - Meaningful only for close observations (for translation at least)

Visual-Inertial SLAM

- Measurements:
  - Acceleration (a_B)
  - Angular velocity (ω_B)
- Needed:
  - Velocity (v_B)
  - Position (p_B)
  - Rotation (R_B)
- Extra parameters
  - Biases (b_a, b_g)
  - Gravity (g_W)
- Euler integration (or Runge-Kutta)
- IMU preintegration



$$\mathbf{R}_{\mathtt{WB}}^{k+1} = \mathbf{R}_{\mathtt{WB}}^{k}\,\mathrm{Exp}\left(\left(\boldsymbol{\omega}_{\mathtt{B}}^{k} - \boldsymbol{b}_{g}^{k}\right)\Delta t\right)$$

$$_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{k+1} = {_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{k}} + \mathbf{g}_{\mathtt{W}}\Delta t + \mathbf{R}_{\mathtt{WB}}^{k}\left(\boldsymbol{a}_{\mathtt{B}}^{k} - \boldsymbol{b}_{a}^{k}\right)\Delta t$$

$$_{\mathtt{W}}\mathbf{p}_{\mathtt{B}}^{k+1} = {_{\mathtt{W}}\mathbf{p}_{\mathtt{B}}^{k}} + {_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{k}}\Delta t + \frac{1}{2}\mathbf{g}_{\mathtt{W}}\Delta t^{2} + \frac{1}{2}\mathbf{R}_{\mathtt{WB}}^{k}\left(\boldsymbol{a}_{\mathtt{B}}^{k} - \boldsymbol{b}_{a}^{k}\right)\Delta t^{2}$$

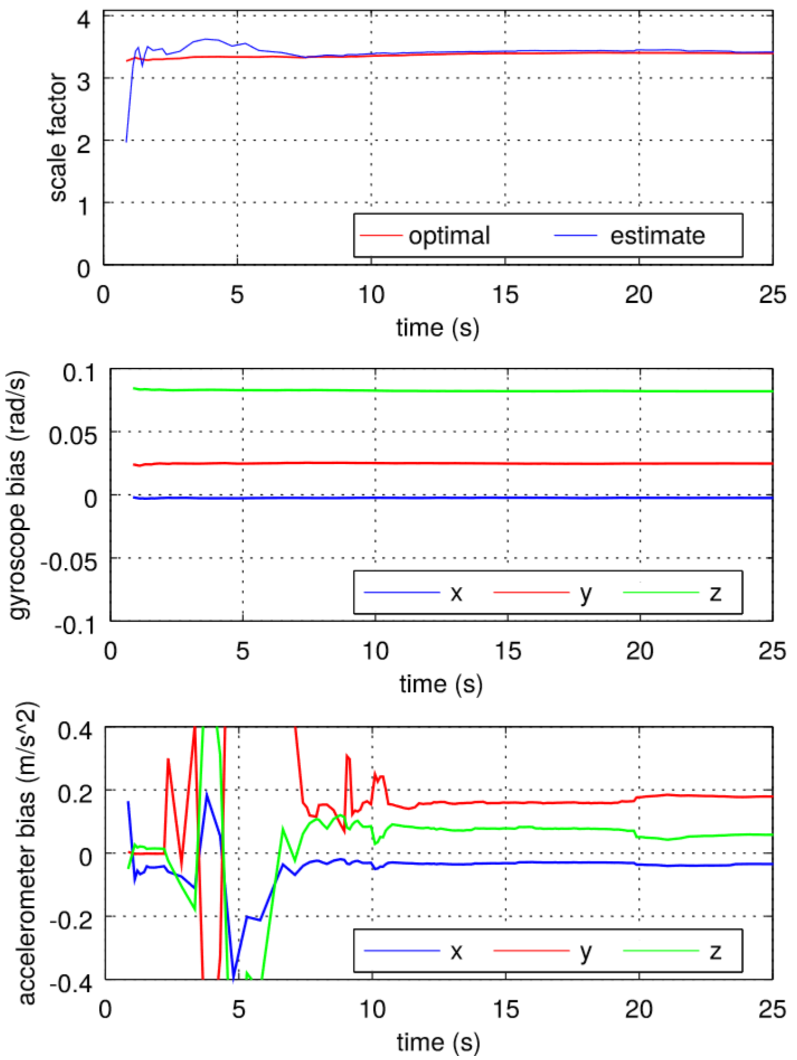
$$\mathbf{R}_{\mathtt{WB}}^{i+1} = \mathbf{R}_{\mathtt{WB}}^{i}\Delta\mathbf{R}_{i,i+1}\mathrm{Exp}\left(\left(\mathbf{J}_{\Delta R}^{g}\mathbf{b}_{g}^{i}\right)\right)$$

$$\begin{aligned} {_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{i+1}} &= {_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{i}} + \mathbf{g}_{\mathtt{W}}\Delta t_{i,i+1} \\ &\quad + \mathbf{R}_{\mathtt{WB}}^{i}\left(\Delta\mathbf{v}_{i,i+1} + \mathbf{J}_{\Delta v}^{g}\mathbf{b}_{g}^{i} + \mathbf{J}_{\Delta v}^{a}\mathbf{b}_{a}^{i}\right) \end{aligned}$$

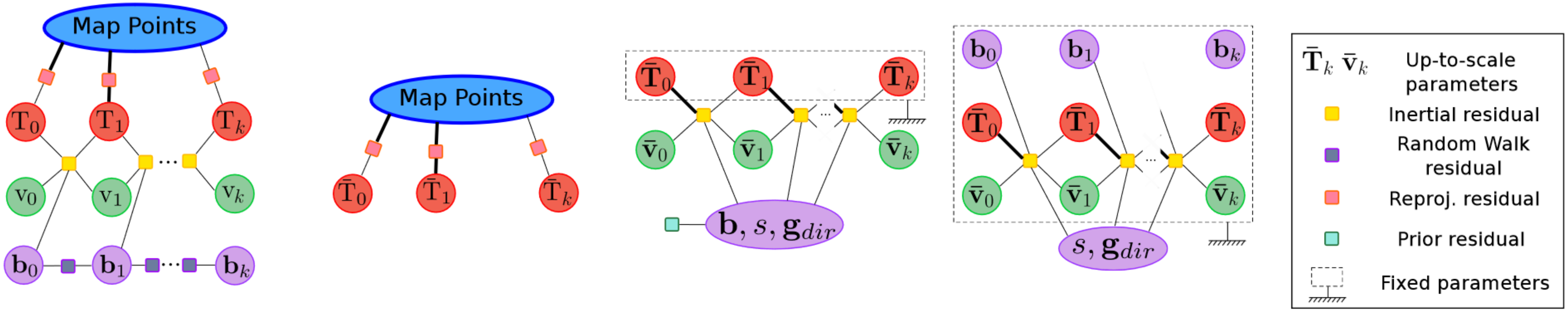
$$\begin{aligned} {_{\mathtt{W}}\mathbf{p}_{\mathtt{B}}^{i+1}} &= {_{\mathtt{W}}\mathbf{p}_{\mathtt{B}}^{i}} + {_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{i}}\Delta t_{i,i+1} + \frac{1}{2}\mathbf{g}_{\mathtt{W}}\Delta t_{i,i+1}^{2} \\ &\quad + \mathbf{R}_{\mathtt{WB}}^{i}\left(\Delta\mathbf{p}_{i,i+1} + \mathbf{J}_{\Delta p}^{g}\mathbf{b}_{g}^{i} + \mathbf{J}_{\Delta p}^{a}\mathbf{b}_{a}^{i}\right) \end{aligned}$$

Visual-Inertial SLAM- Calibration

- IMU initialization
- Gyroscope bias estimation
- Scale and gravity estimation
- Accelerometer Bias Estimation
- Scale and Gravity Direction Refinement

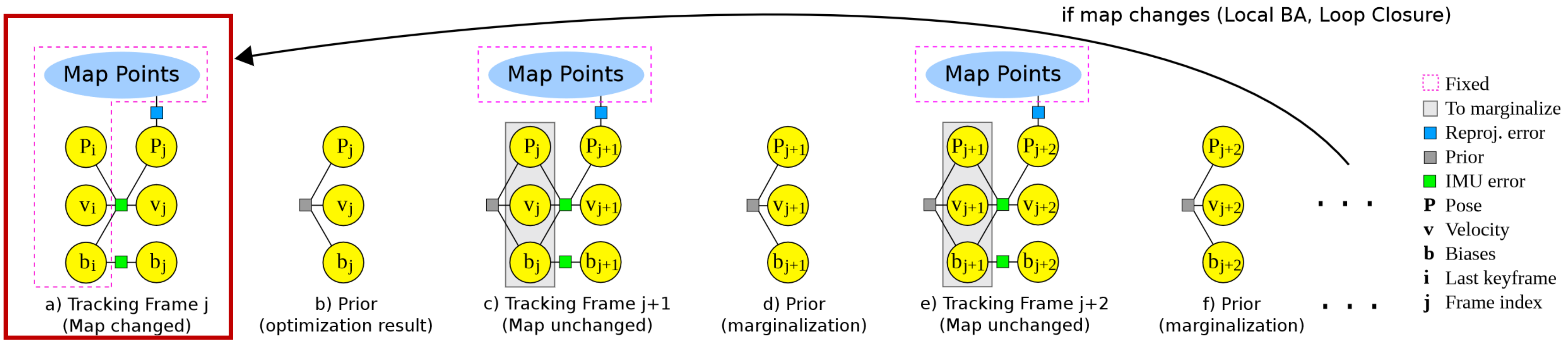


Visual-Inertial SLAM- ORB-SLAM 3



(a) Visual-Inertial (b) Visual-Only (c) Inertial-Only (d) Scale and Gravity

Visual-Inertial SLAM- Tracking



$$\theta = \left\{ \mathbf{R}^j_{\mathtt{WB}}, {}_{\mathtt{W}}\mathbf{p}^j_{\mathtt{B}}, {}_{\mathtt{W}}\mathbf{v}^j_{\mathtt{B}}, \mathbf{b}^j_g, \mathbf{b}^j_a \right\}$$

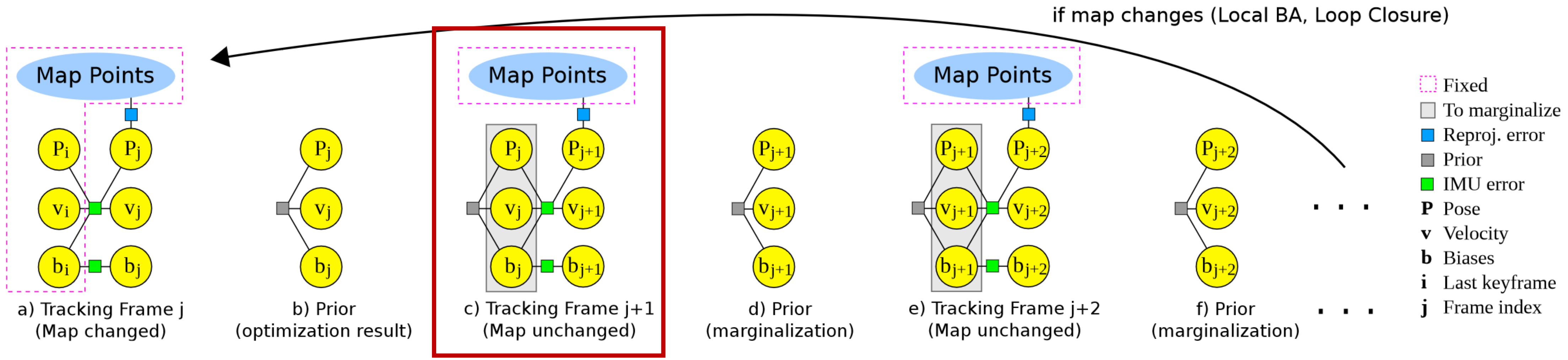
$$\theta^* = \underset{\theta}{\operatorname{argmin}} \left(\sum_k \mathbf{E}_{\text{proj}}(k,j) + \mathbf{E}_{\text{IMU}}(i,j) \right)$$

$$\mathbf{E}_{\text{proj}}(k,j) = \rho\left(\left(\mathbf{x}^k - \pi(\mathbf{X}^k_{\mathtt{C}})\right)^T \boldsymbol{\Sigma}_{\boldsymbol{k}} \left(\mathbf{x}^k - \pi(\mathbf{X}^k_{\mathtt{C}})\right) \right)$$

$$\mathbf{X}^k_{\mathtt{C}} = \mathbf{R}_{\mathtt{CB}} \mathbf{R}^j_{\mathtt{BW}} \left(\mathbf{X}^k_{\mathtt{W}} - {}_{\mathtt{W}}\mathbf{p}^j_{\mathtt{B}} \right) + {}_{\mathtt{C}}\mathbf{p}_{\mathtt{B}}$$

$$\begin{aligned}
\mathbf{E}_{\text{IMU}}(i,j) &= \rho\left(\left[\mathbf{e}_R^T\, \mathbf{e}_v^T\, \mathbf{e}_p^T\right] \boldsymbol{\Sigma}_I \left[\mathbf{e}_R^T\, \mathbf{e}_v^T\, \mathbf{e}_p^T\right]^T \right) \\
&\quad + \rho\left(\mathbf{e}_b^T \boldsymbol{\Sigma}_R \mathbf{e}_b \right) \\
\mathbf{e}_R &= \operatorname{Log}\left(\left(\Delta\mathbf{R}_{ij} \operatorname{Exp}\left(\mathbf{J}^g_{\Delta R} \mathbf{b}^j_g \right) \right)^T \mathbf{R}^i_{\mathtt{BW}} \mathbf{R}^j_{\mathtt{WB}} \right) \\
\mathbf{e}_v &= \mathbf{R}^i_{\mathtt{BW}} \left({}_{\mathtt{W}}\mathbf{v}^j_{\mathtt{B}} - {}_{\mathtt{W}}\mathbf{v}^i_{\mathtt{B}} - \mathbf{g}_{\mathtt{W}} \Delta t_{ij} \right) \\
&\quad - \left(\Delta\mathbf{v}_{ij} + \mathbf{J}^g_{\Delta v} \mathbf{b}^j_g + \mathbf{J}^a_{\Delta v} \mathbf{b}^j_a \right) \\
\mathbf{e}_p &= \mathbf{R}^i_{\mathtt{BW}} \left({}_{\mathtt{W}}\mathbf{p}^j_{\mathtt{B}} - {}_{\mathtt{W}}\mathbf{p}^i_{\mathtt{B}} - {}_{\mathtt{W}}\mathbf{v}^i_{\mathtt{B}} \Delta t_{ij} - \frac{1}{2} \mathbf{g}_{\mathtt{W}} \Delta t^2_{ij} \right) \\
&\quad - \left(\Delta\mathbf{p}_{ij} + \mathbf{J}^g_{\Delta p} \mathbf{b}^j_g + \mathbf{J}^a_{\Delta p} \mathbf{b}^j_a \right) \\
\mathbf{e}_b &= \mathbf{b}^j - \mathbf{b}^i
\end{aligned}$$

Visual-Inertial SLAM- Tracking



$$\theta = \left\{ \mathbf{R}_{\mathtt{WB}}^{j}, \mathbf{p}_{\mathtt{W}}^{j}, \mathbf{v}_{\mathtt{W}}^{j}, \mathbf{b}_{g}^{j}, \mathbf{b}_{a}^{j}, \mathbf{R}_{\mathtt{WB}}^{j+1}, \mathbf{p}_{\mathtt{W}}^{j+1}, \mathbf{v}_{\mathtt{W}}^{j+1}, \mathbf{b}_{g}^{j+1}, \mathbf{b}_{a}^{j+1} \right\}$$

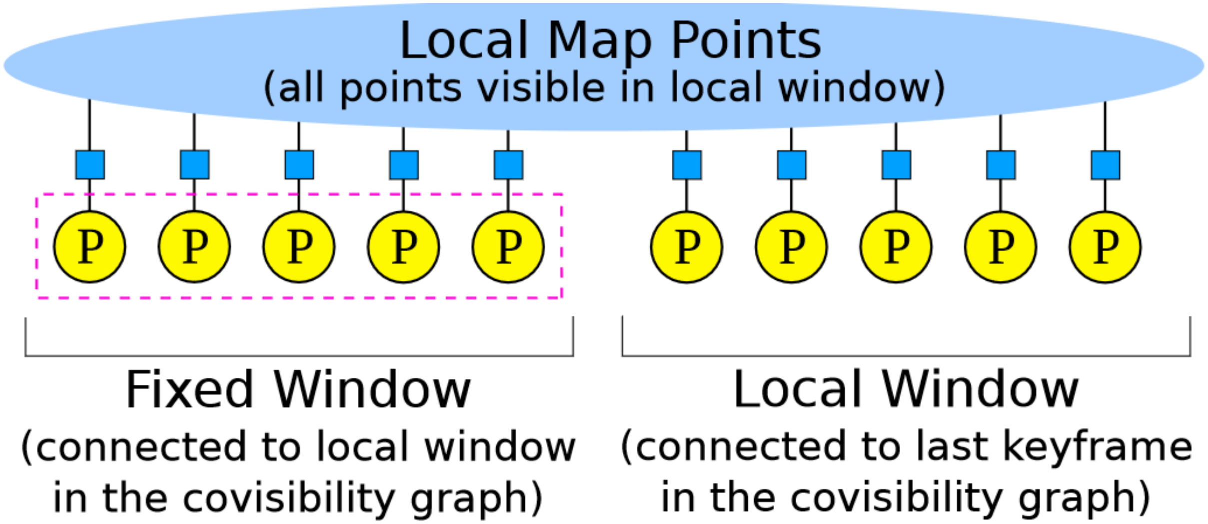
$$\theta^{*} = \underset{\theta}{\operatorname{argmin}} \left(\sum_{k} \mathbf{E}_{\text{proj}}(k, j+1) + \mathbf{E}_{\text{IMU}}(j, j+1) + \mathbf{E}_{\text{prior}}(j) \right)$$

$$\mathbf{E}_{\text{prior}}(j) = \rho \left(\left[\mathbf{e}_R^T \, \mathbf{e}_v^T \, \mathbf{e}_p^T \, \mathbf{e}_b^T \right] \boldsymbol{\Sigma}_p \left[\mathbf{e}_R^T \, \mathbf{e}_v^T \, \mathbf{e}_p^T \, \mathbf{e}_b^T \right]^T \right)$$

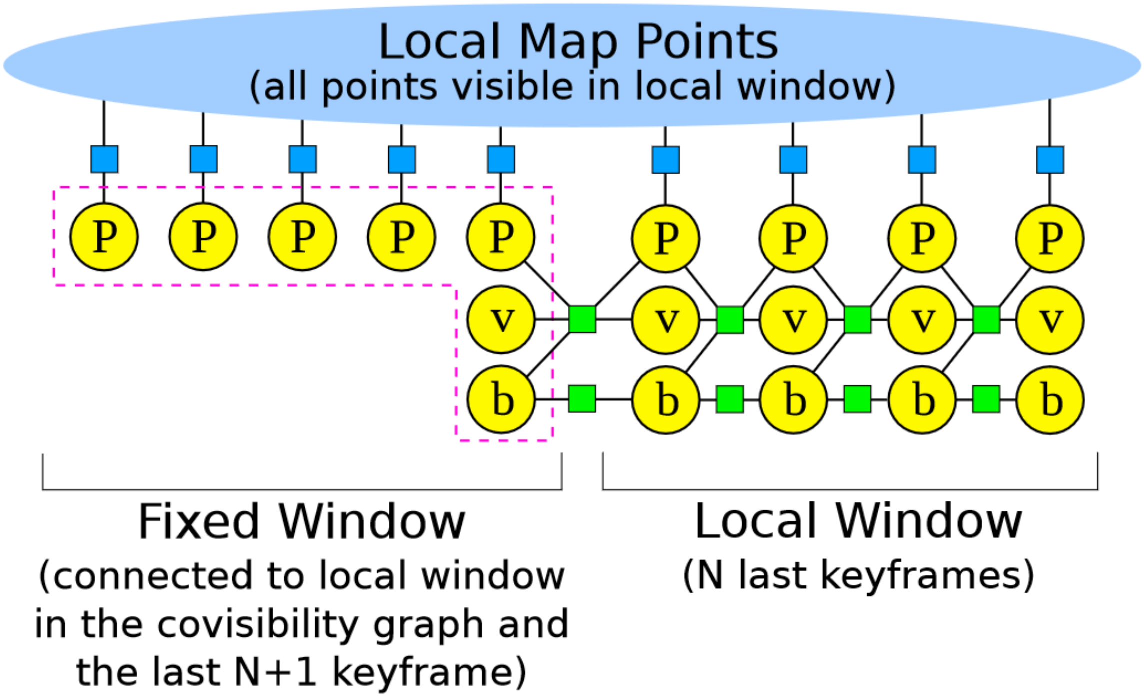
$$\mathbf{e}_R = \operatorname{Log}\left(\bar{\mathbf{R}}_{\mathtt{BW}}^{j} \mathbf{R}_{\mathtt{WB}}^{j} \right) \qquad \mathbf{e}_v = {}_{\mathtt{W}}\bar{\mathbf{v}}_{\mathtt{B}}^{j} - {}_{\mathtt{W}}\mathbf{v}_{\mathtt{B}}^{j}$$

$$\mathbf{e}_p = {}_{\mathtt{W}}\bar{\mathbf{p}}_{\mathtt{B}}^{j} - {}_{\mathtt{W}}\mathbf{p}_{\mathtt{B}}^{j} \qquad \mathbf{e}_b = \bar{\mathbf{b}}^{j} - \mathbf{b}^{j}$$

Visual-Inertial SLAM – Local BA



ORB-SLAM's Local BA



Visual-Inertial ORB-SLAM's Local BA

	SLAM or VO	Pixels used	Data association	Estimation	Relocalization	Loop closing	Multi Maps	Mono	Stereo	Mono IMU	Stereo IMU	Fisheye	Accuracy	Robustness	Open source
Mono-SLAM [13], [14]	SLAM	Shi Tomasi	Correlation	EKF	-	-	-	✓	-	-	-	-	Fair	Fair	[15][1]
PTAM [16]–[18]	SLAM	FAST	Pyramid SSD	BA	Thumbnail	-	-	✓	-	-	-	-	Very Good	Fair	[19]
LSD-SLAM [20], [21]	SLAM	Edgelets	Direct	PG	-	FABMAP PG	-	✓	✓	-	-	-	Good	Fair	[22]
SVO [23], [24]	VO	FAST+ Hi.grad.	Direct	Local BA	-	-	-	✓	✓	-	-	✓	Very Good	Very Good	[25][2]
ORB-SLAM2 [2], [3]	SLAM	ORB	Descriptor	Local BA	DBoW2	DBoW2 PG+BA	-	✓	✓	-	-	-	Exc.	Very Good	[26]
DSO [27]–[29]	VO	High grad.	Direct	Local BA	-	-	-	✓	✓	-	-	✓	Fair	Very Good	[30]
DSM [31]	SLAM	High grad.	Direct	Local BA	-	-	-	✓	-	-	-	-	Very Good	Very Good	[32]
MSCKF [33]–[36]	VO	Shi Tomasi	Cross correlation	EKF	-	-	-	✓	-	✓	✓	-	Fair	Very Good	[37][3]
OKVIS [38], [39]	VO	BRISK	Descriptor	Local BA	-	-	-	-	-	✓	✓	✓	Good	Very Good	[40]
ROVIO [41], [42]	VO	Shi Tomasi	Direct	EKF	-	-	-	-	-	✓	✓	✓	Good	Very Good	[43]
ORBSLAM-VI [4]	SLAM	ORB	Descriptor	Local BA	DBoW2	DBoW2 PG+BA	-	✓	-	✓	-	-	Very Good	Very Good	-
VINS-Fusion [7], [44]	VO	Shi Tomasi	KLT	Local BA	DBoW2	DBoW2 PG	✓	-	✓	✓	✓	✓	Good	Exc.	[45]
VI-DSO [46]	VO	High grad.	Direct	Local BA	-	-	-	-	-	✓	-	-	Very Good	Exc.	-
BASALT [47]	VO	FAST	KLT (LSSD)	Local BA	-	ORB BA	-	-	-	-	✓	✓	Very Good	Exc.	[48]
Kimera [8]	VO	Shi Tomasi	KLT	Local BA	-	DBoW2 PG	-	-	-	-	✓	-	Good	Exc.	[49]
ORB-SLAM3 (ours)	SLAM	ORB	Descriptor	Local BA	DBoW2	DBoW2 PG+BA	✓	✓	✓	✓	✓	✓	Exc.	Exc.	[5]

Out of scope

- Dense and direct approaches
- Deep learning-based solutions
- Multi map SLAM
- Image retrieval
- Map initialization

Sources

1. Engel, J., Koltun, V., & Cremers, D. (2016). Direct Sparse Odometry. *ArXiv*. /abs/1607.02565
2. Montiel, J. M., & Tardos, J. D. (2015). ORB-SLAM: A Versatile and Accurate Monocular SLAM System. *ArXiv*. https://doi.org/10.1109/TRO.2015.2463671
3. Tardos, J. D. (2016). Visual-Inertial Monocular SLAM with Map Reuse. *ArXiv*. https://doi.org/10.1109/LRA.2017.2653359
4. Tardos, J. D. (2016). ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras. *ArXiv*. https://doi.org/10.1109/TRO.2017.2705103
5. Campos, C., Elvira, R., Rodríguez, J. J., Montiel, J. M., & Tardós, J. D. (2020). ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. *ArXiv*. https://doi.org/10.1109/TRO.2021.3075644
6. Höll, M., & Lepetit, V. (2017). Monocular LSD-SLAM Integration within AR System. *ArXiv*. https://doi.org/10.13140/RG.2.2.10054.27205
7. E. Rublee, V. Rabaud, K. Konolige and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," *2011 International Conference on Computer Vision*, Barcelona, Spain, 2011, pp. 2564-2571, doi: 10.1109/ICCV.2011.6126544.
8. Calonder, M., Lepetit, V., Strecha, C., Fua, P. (2010). BRIEF: Binary Robust Independent Elementary Features. In: Daniilidis, K., Maragos, P., Paragios, N. (eds) Computer Vision – ECCV 2010. ECCV 2010. Lecture Notes in Computer Science, vol 6314. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-642-15561-1_56
9. Wikimedia Foundation. (2023, August 26). *Features from accelerated segment test*. Wikipedia. https://en.wikipedia.org/wiki/Features_from_accelerated_segment_test
10. Forster, C., Carlone, L., Dellaert, F., & Scaramuzza, D. (2015). On-Manifold Preintegration for Real-Time Visual-Inertial Odometry. *ArXiv*. https://doi.org/10.1109/TRO.2016.2597321