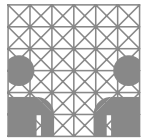


Kimera: Semantic SLAM

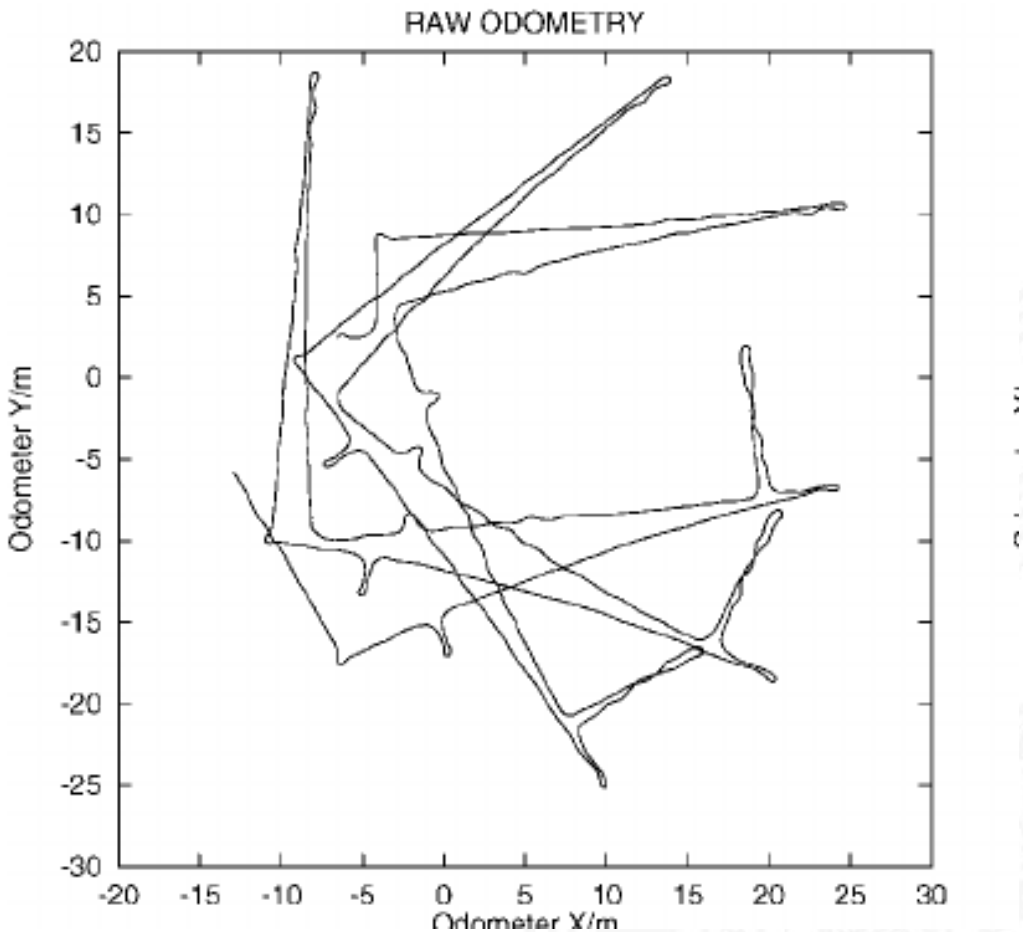
University of Hamburg
Faculty of Mathematics, Informatics and Natural Sciences
Department of Informatics

Technical Aspects of Multimodal Systems

28. May 2020

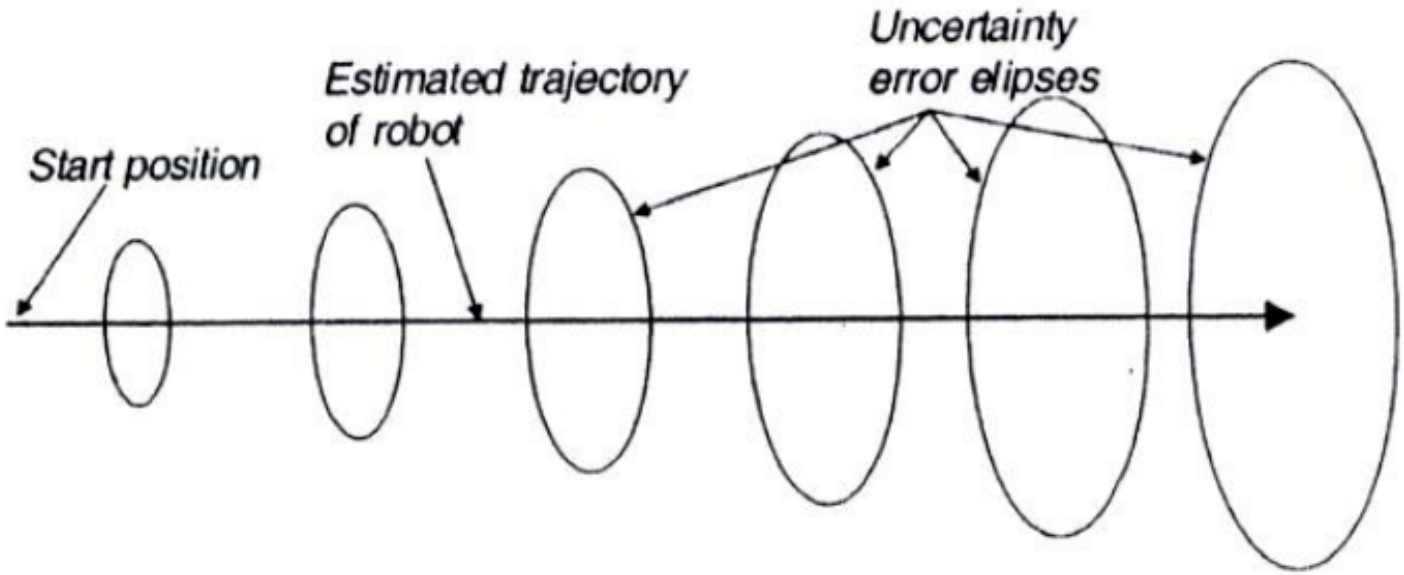
Outline

Motivation



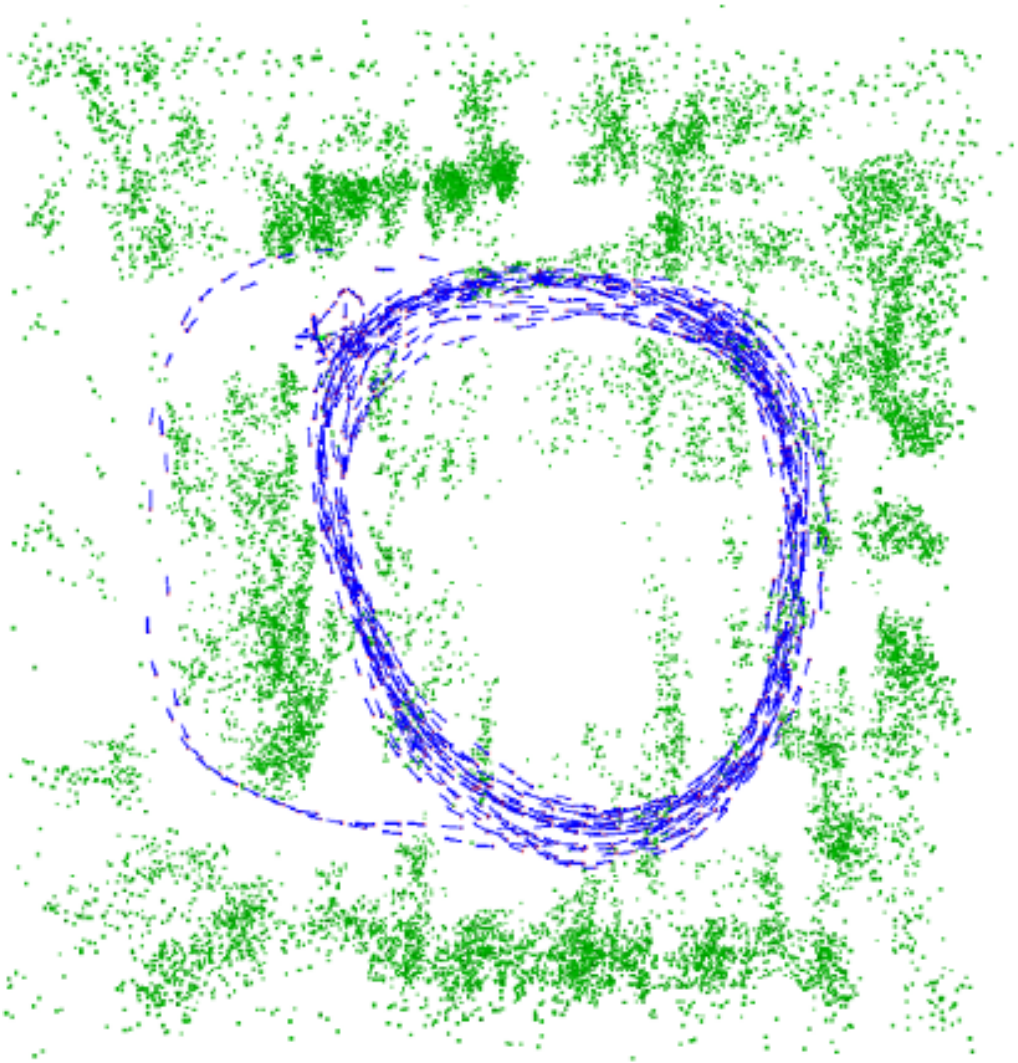
Source: [Amer et al., 2012]

Motivation



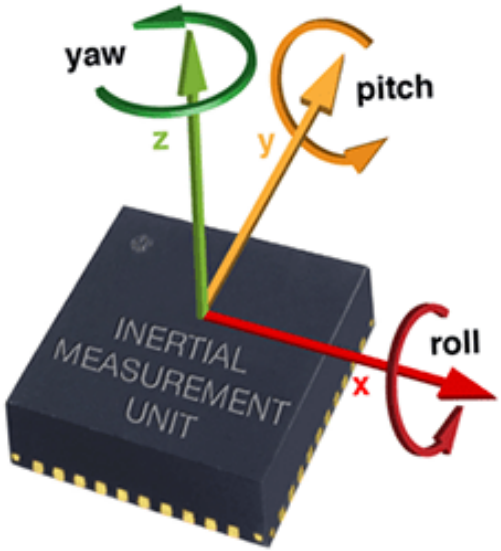
Source: [Amer et al., 2012]

Motivation



Source: [Forster et al., 2015]

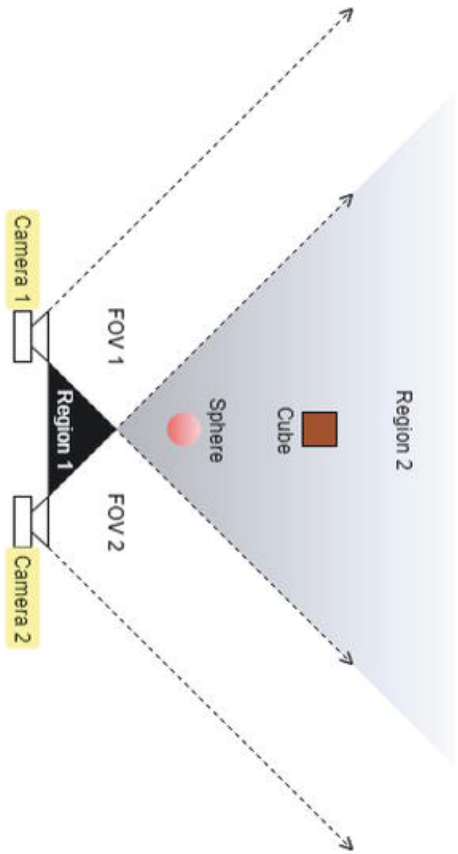
Inertial Measurement Unit



Source: https://vrtracker.xyz/handling-imu-drift/

IMU

- Components:
  - Accelerometer
  - Gyroscope
  - Magnetometer
- Acceleration and Rotation are measured for all three euclidian axis

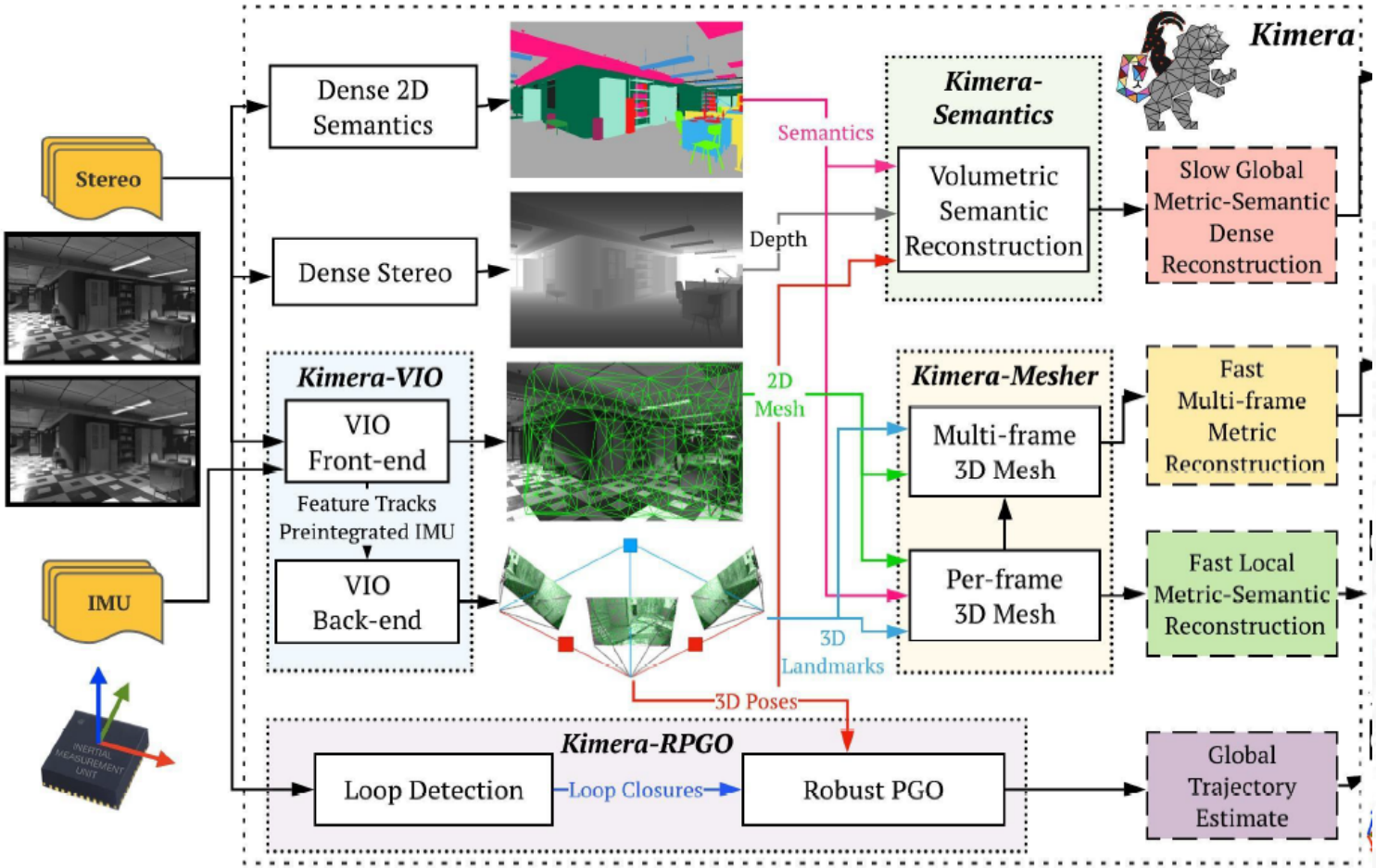


Source: https://www.intechopen.com/books/coding-theory/efficient-depth-estimation-using-sparse-stereo-vision-with-other-perception-techniques

Stereo Camera

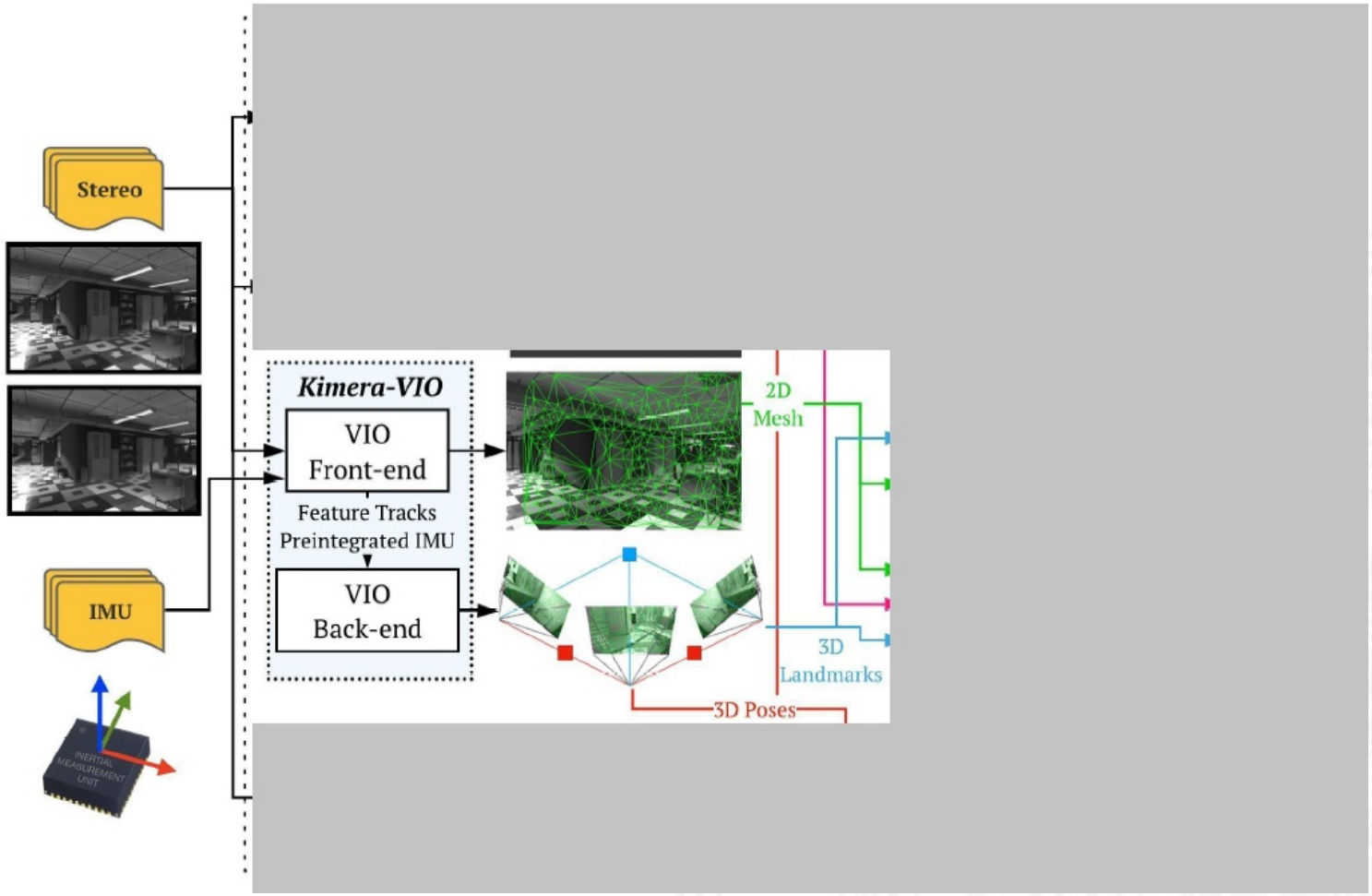
- 2 aligned monocular cameras
- Distance (baseline) between the cameras is known
- The distance between an object, located in the overlapping parts of the FOVs and the cameras can then be calculated

Overview



Source: [Rosinol et al., 2019a]

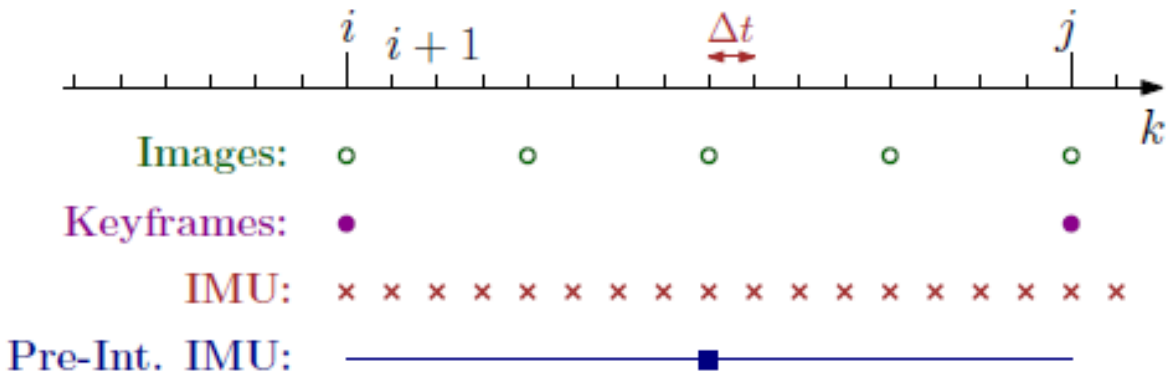
Visual Inertial Odometry (VIO)



Source: [Rosinol et al., 2019a]

Manifold Preintegration for IMUs

- Idea:
  - Take the raw IMU measurements between two keyframes
  - Calculate a motion constraint (Preintegrated IMU Measurement) between both keyframes

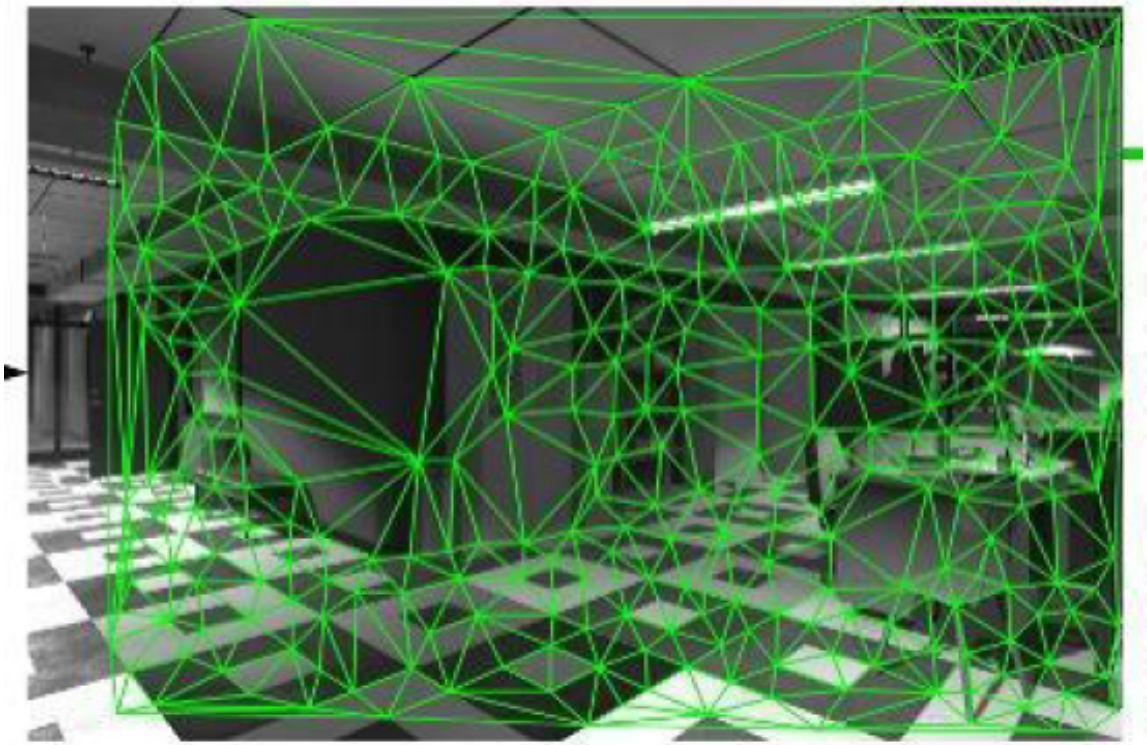


Source: [Forster et al., 2015]

Vision Front-End

- For Keyframes:
  - 1. Stereo Matching
  - 2. Identify Landmarks (Features)
    - Shi-Tomasi Corner Detection
    - Detect Corners via threshold $R = min(\lambda_1, \lambda_2)$
  - 3. Geometric Verification
    - Different variations of RANSAC-Algorithm
- For Intermediate Frames:
  - Track Landmarks
    - Lucas-Kanade Tracker
    - Observe a neighbourhood of Voxels in an image
    - Calculate the movement based on the differences in the images

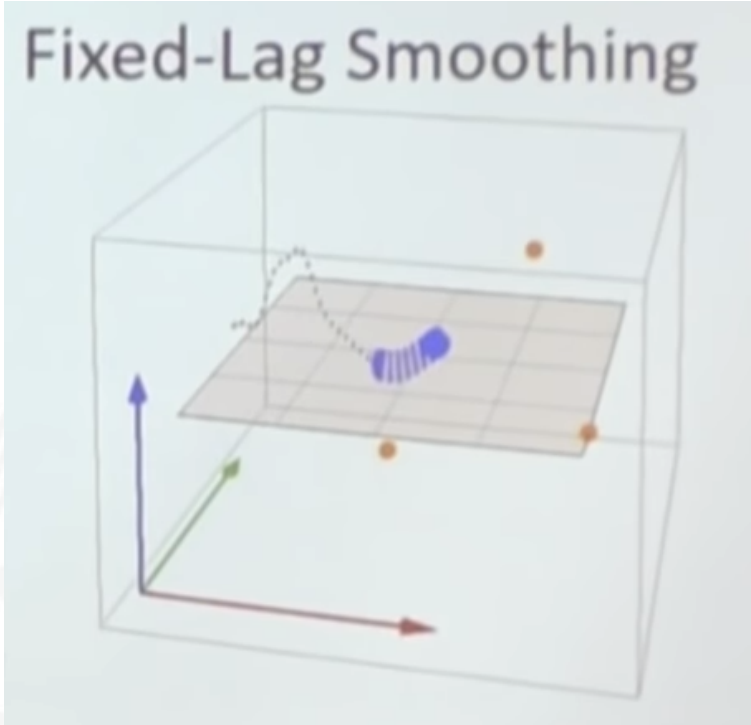
Feature/Landmark Detection



Source: [Rosinol et al., 2019a]

VIO-Backend - Fixed Lag Smoothing

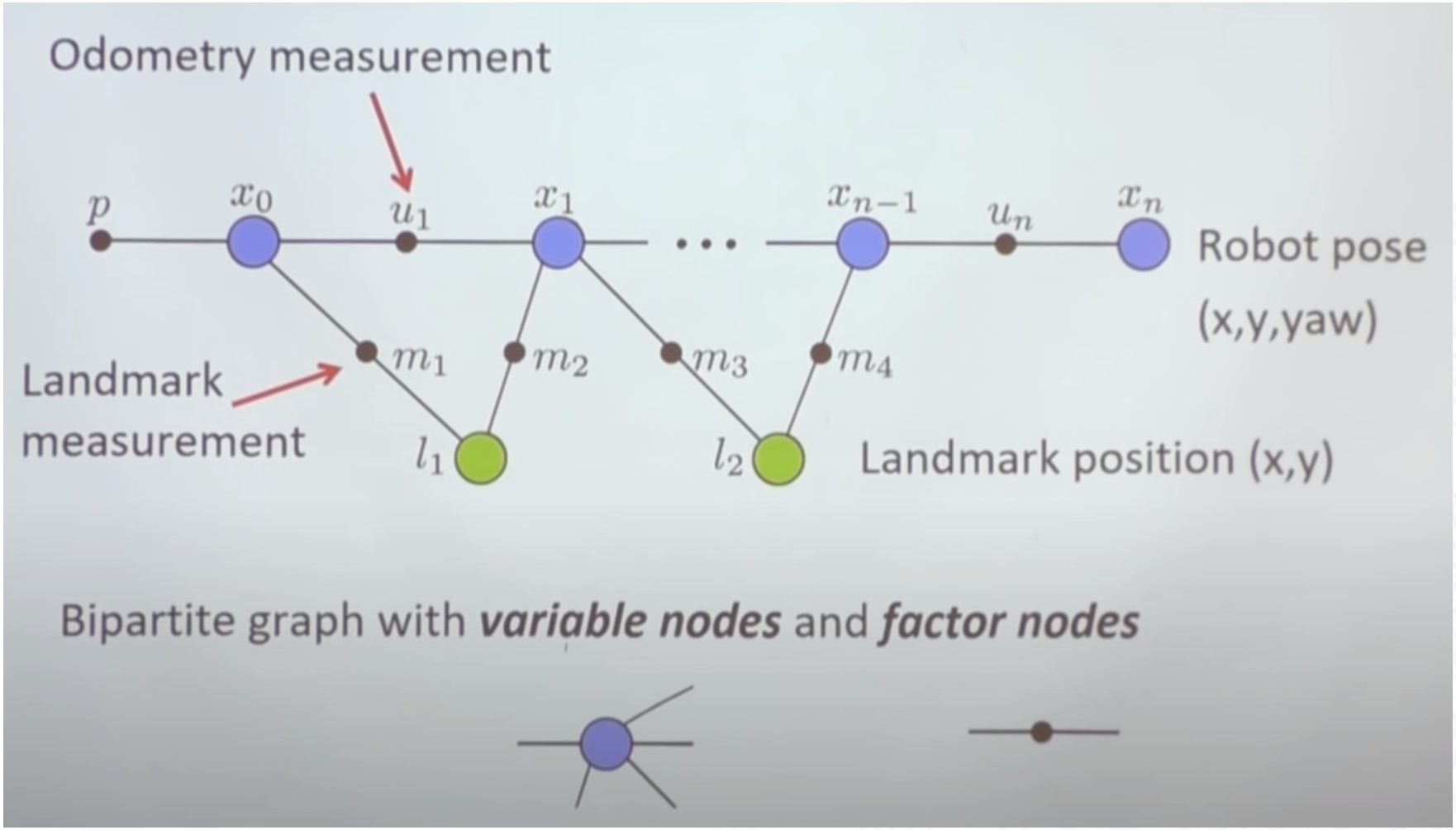
- Fixed-lag Smoothing:
  - Estimate states in a given time window
  - Marginalize older states
- Advantages:
  - more accurate than filtering (e.g. EKF), for nonlinear problems
  - Relinearization of past observations
- Disadvantages:
  - Marginalization lowers accuracy
  - Common with Filtering
    - Consistency is problematic
    - Linearization errors



Source:

https://www.youtube.com/watch?v=Q313pTMAdcM

Factor Graphs



Source: https://www.youtube.com/watch?v=Q313pTMAdcM

Factor Graphs

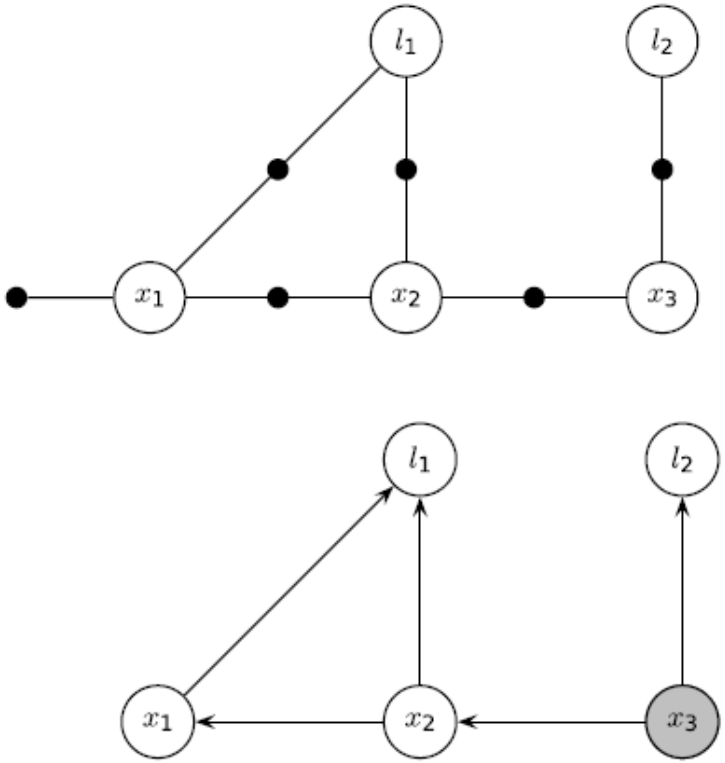
- How do we predict an pose at time t_i from the factor graph?
  - 1. Convert the factor graph to a bayes network
  - 2. Convert bayes network to bayes tree
  - 3. Solve it with iSAM2-Algorithm
    - Part of GTSAM, C++ Open-Source Library
- But why?
  - The bayes tree can be updated incremetally
  - **Which also means** that we only need to calculate the updated parts of the tree
  - This makes it possible to apply it for real-time odometry
- Output

$$argmin_\theta = \sum_i ||h_i(\theta) - z_i||_\theta{}^2 \tag{1}$$

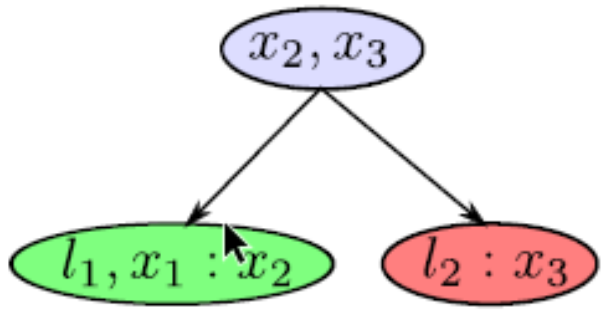
- Linearized:

$$argmin_\theta ||A\theta - b||^2 \tag{2}$$

Bayesian Network & Tree



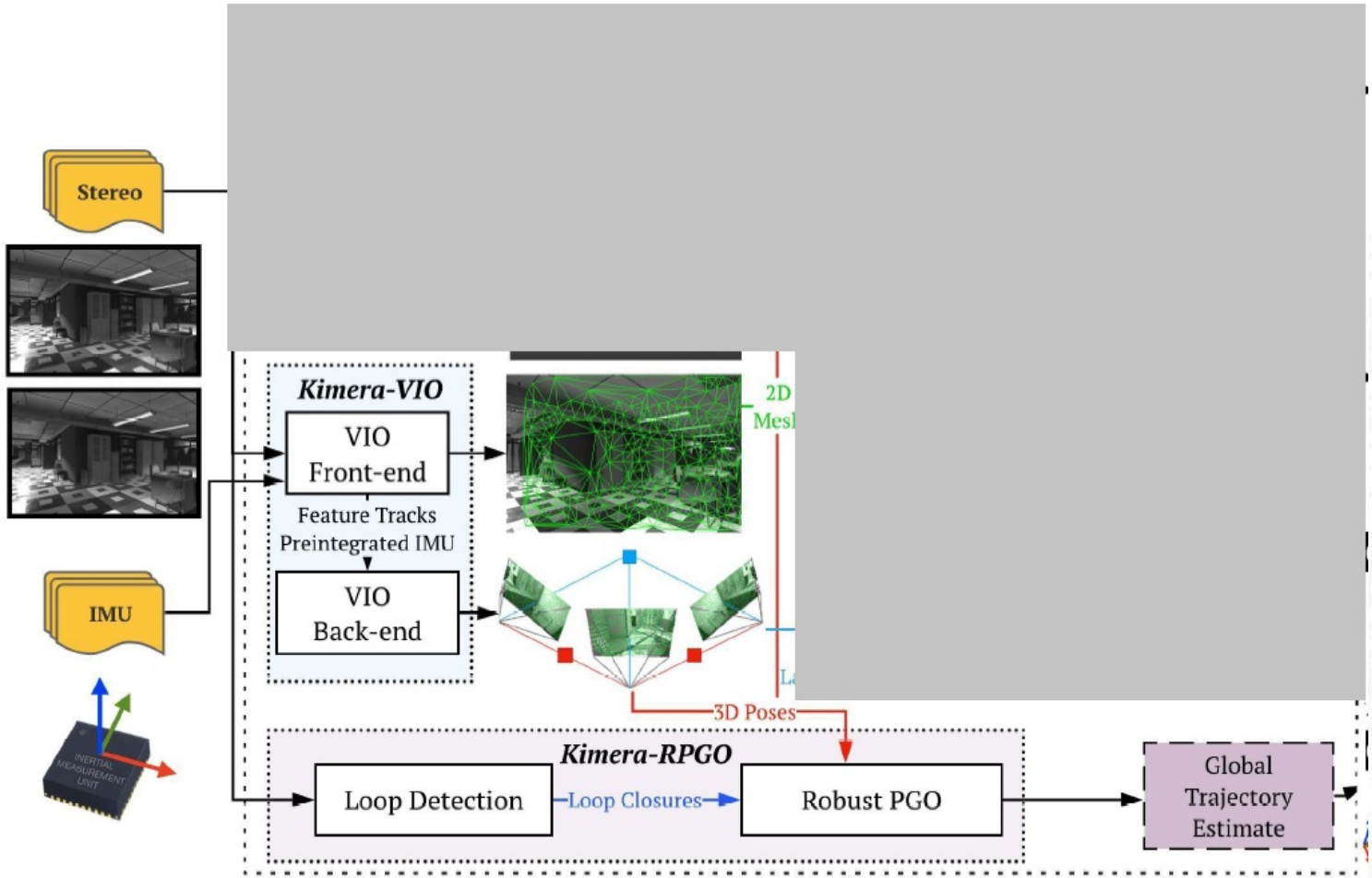
Source: [Kaess et al., 2010]



Source: [Kaess et al., 2010]

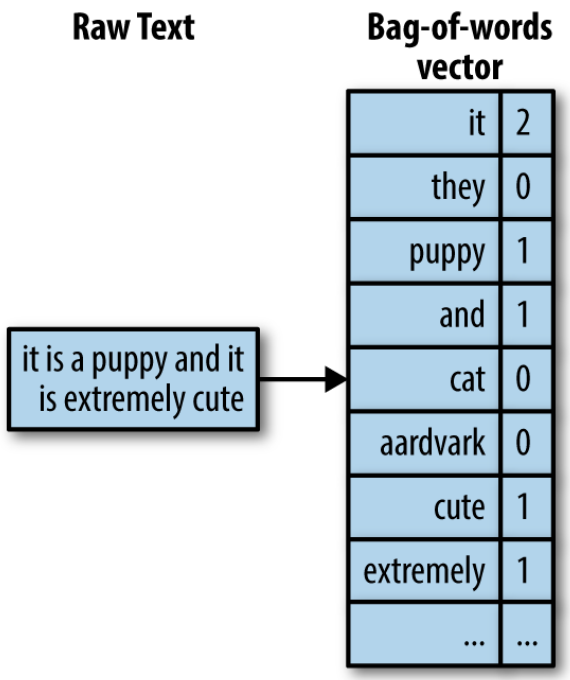
Video

Robust Pose Graph Optimization (RPGO)

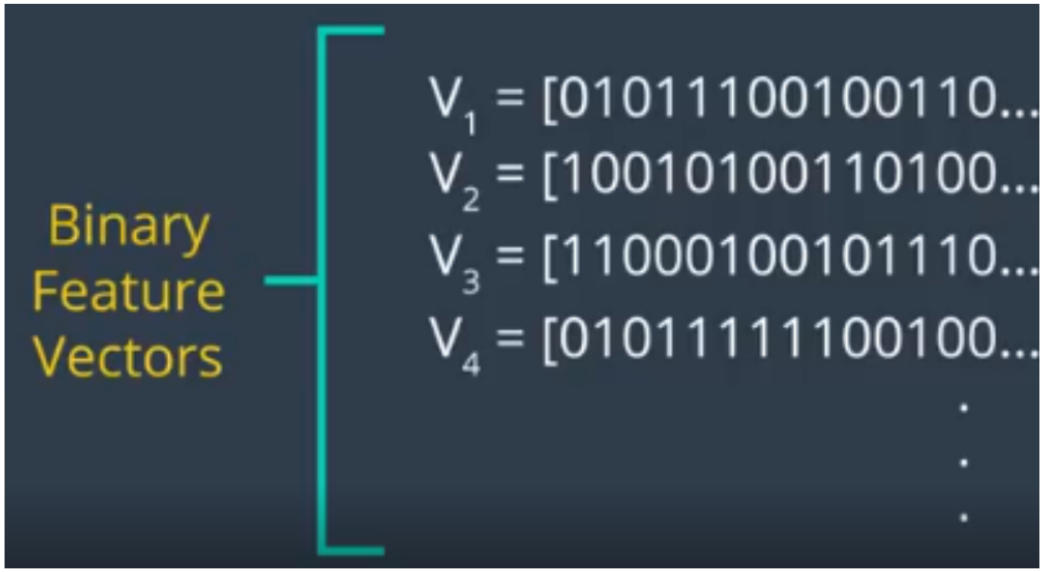


Source: [Rosinol et al., 2019a]

Loop Closures - Bag of Words

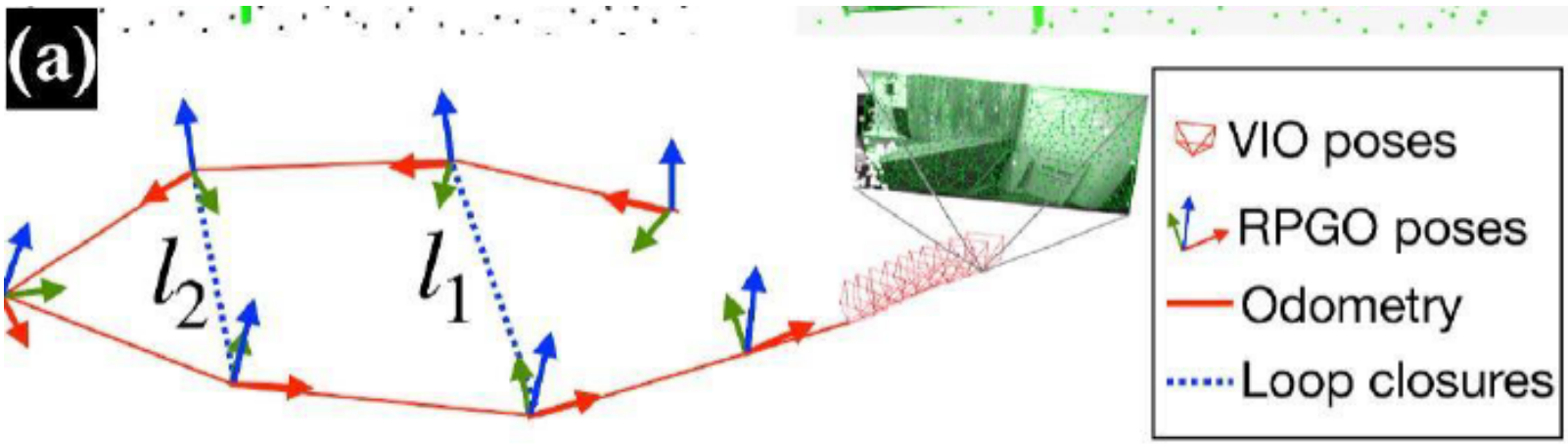


Source: http://uc-r.github.io/creating-text-features



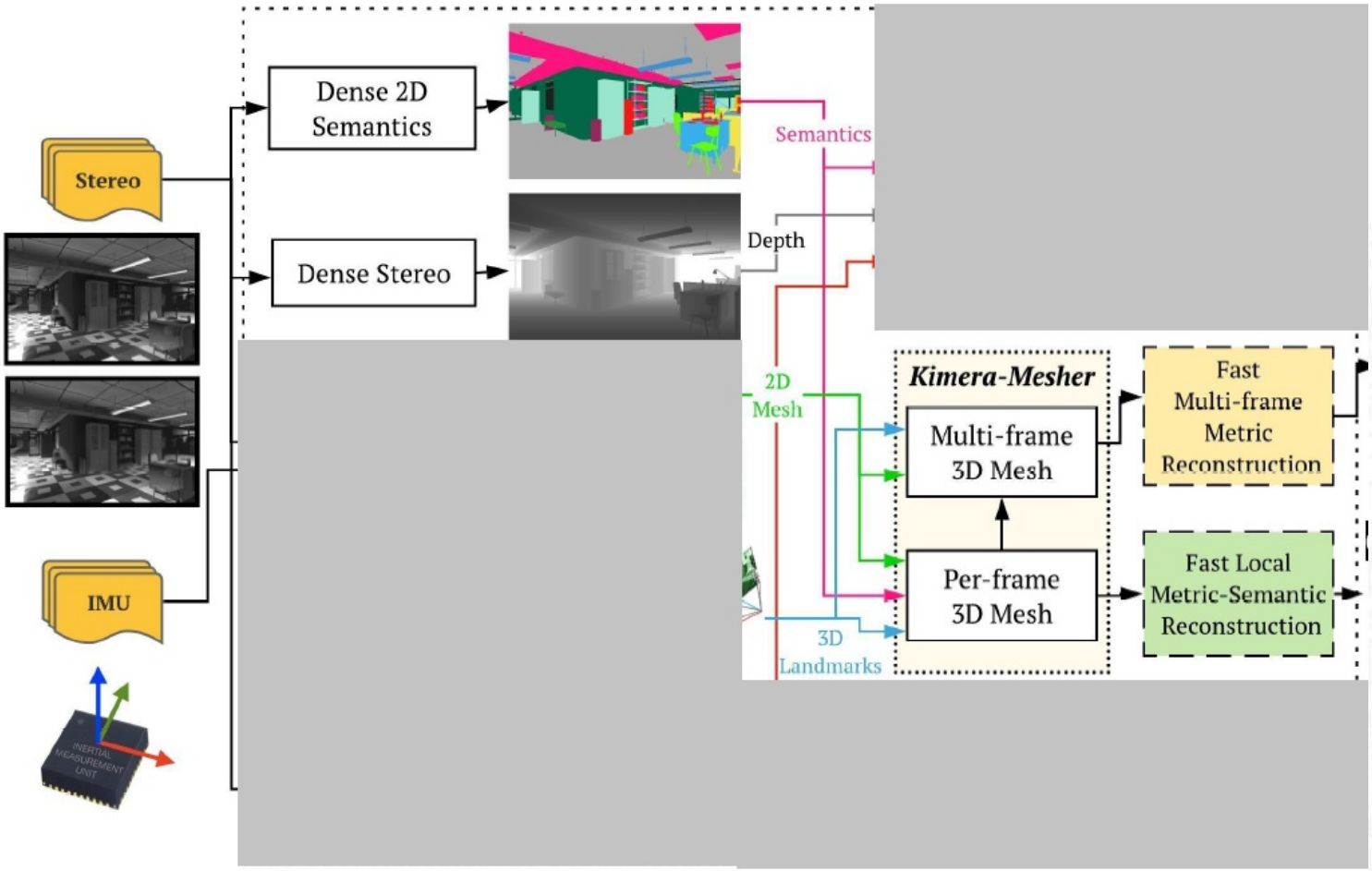
Source: https://medium.com/data-breach/introduction-to-brief-binary-robust-independent-elementary-features-436f4a31a0e6

Global Trajectory Estimate



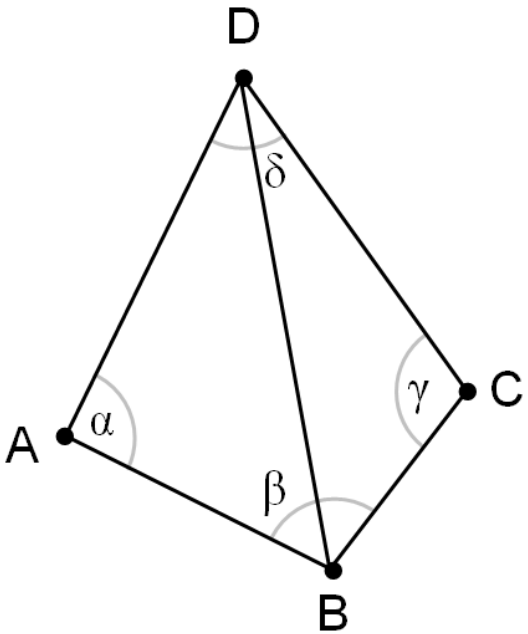
Source: [Rosinol et al., 2019a]

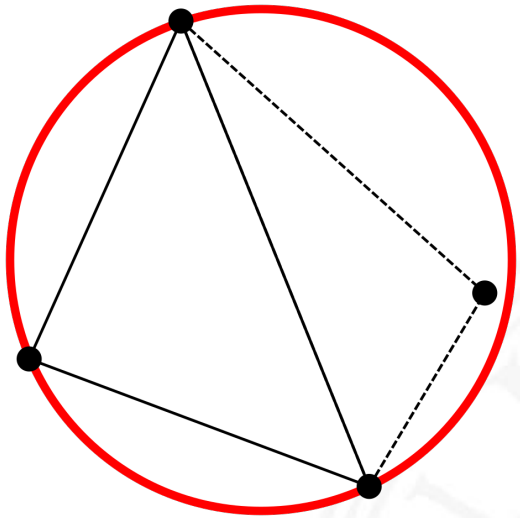
3D Mesh Reconstruction

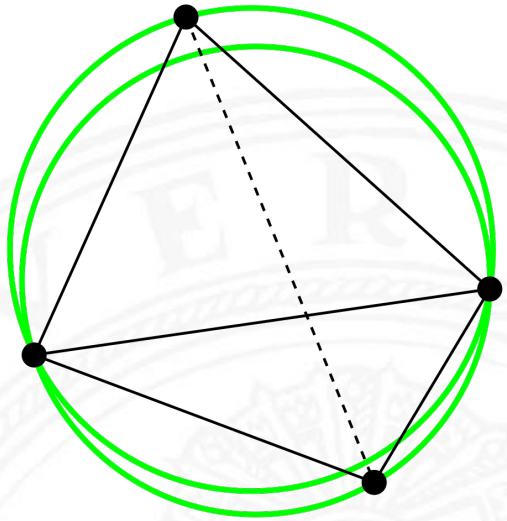


Source: [Rosinol et al., 2019a]

2D Mesh - Delaunay Triangulation



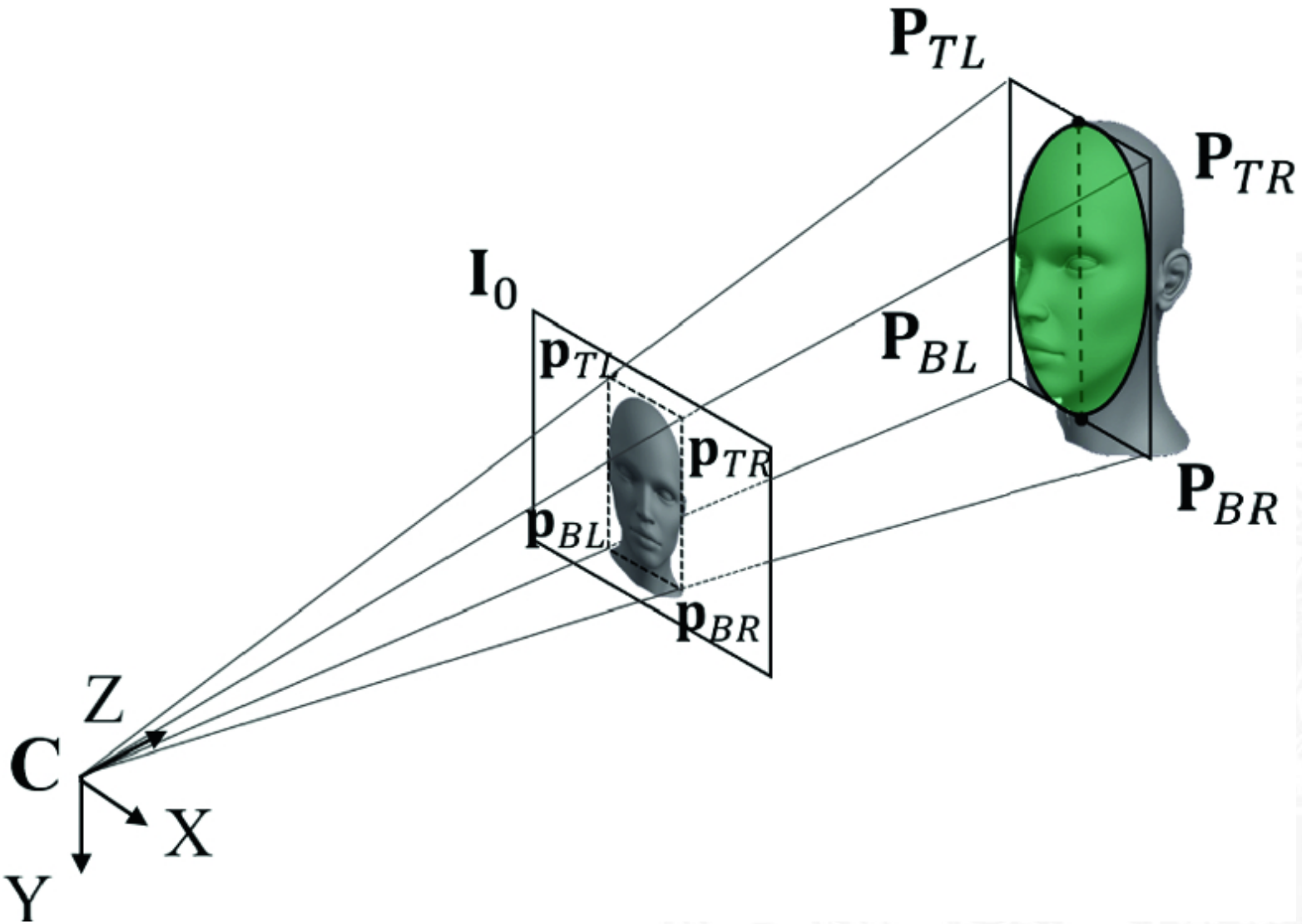




Source:

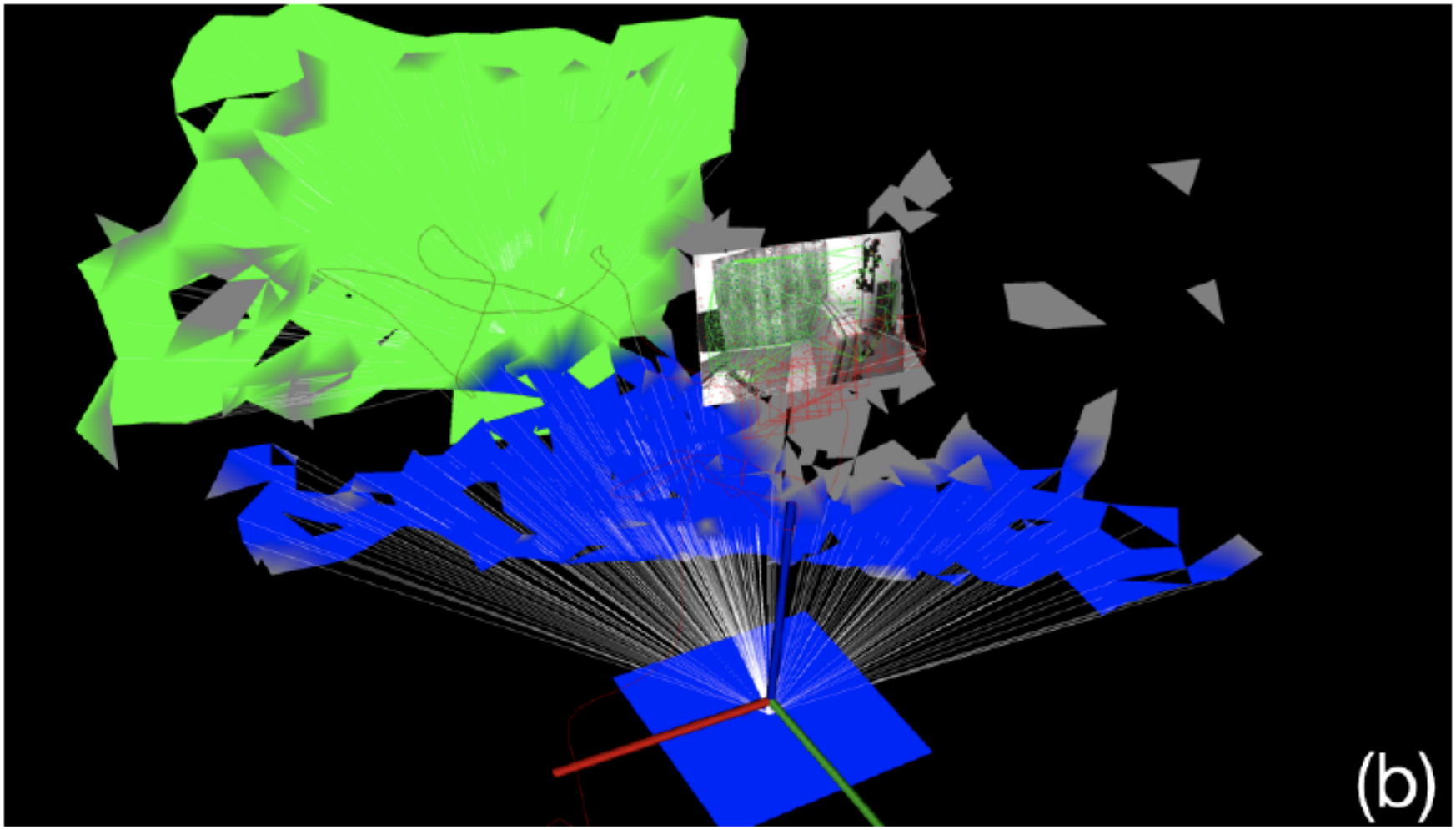
https://en.wikipedia.org/wiki/Delaunay_triangulation

3D Mesh - 2D Mesh Backprojection



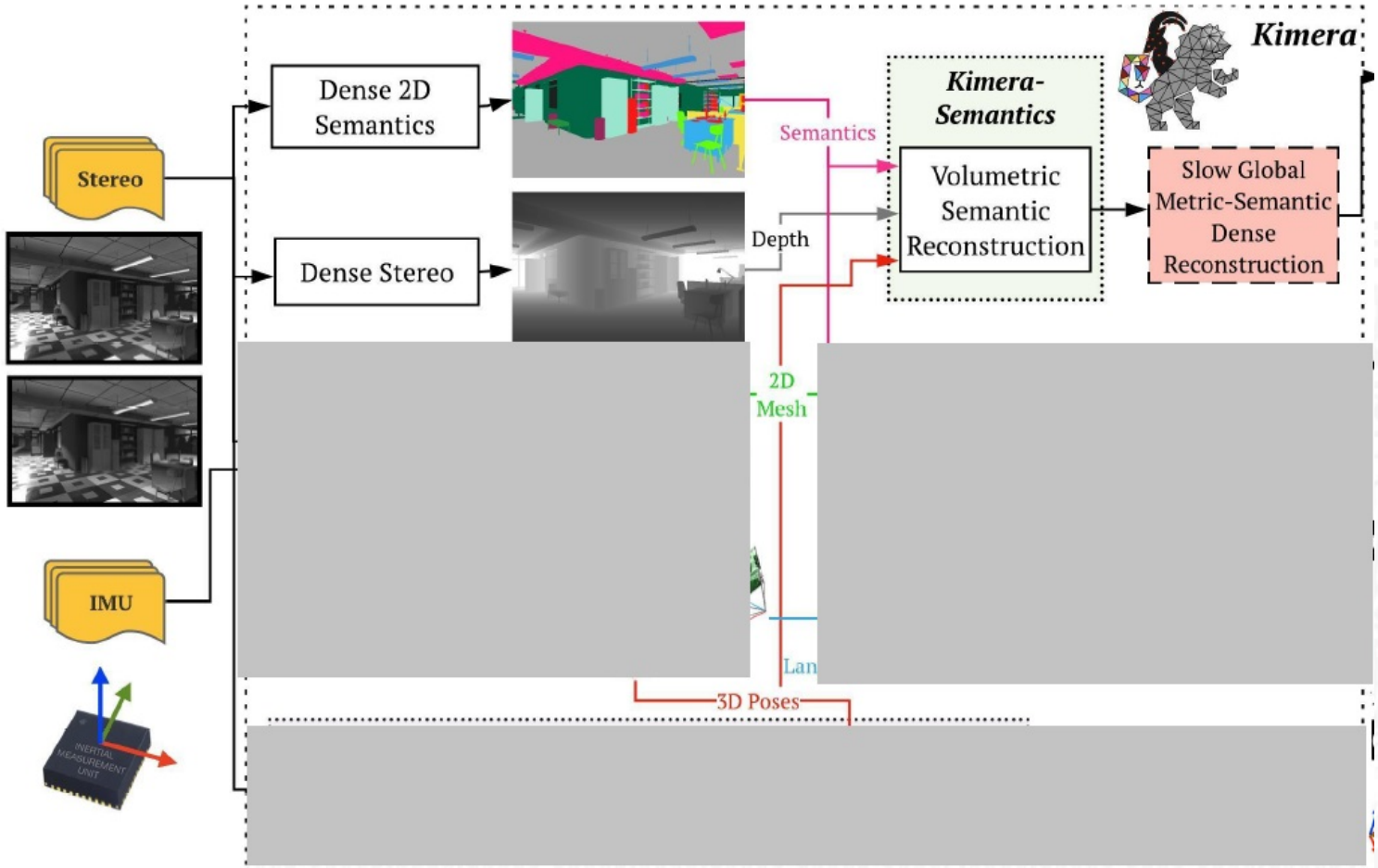
Source: [Díaz Barros et al., 2019]

Multi-Frame 3D Mesh - Mesh Fusion



Source: [Rosinol et al., 2019b]

Kimera Semantics



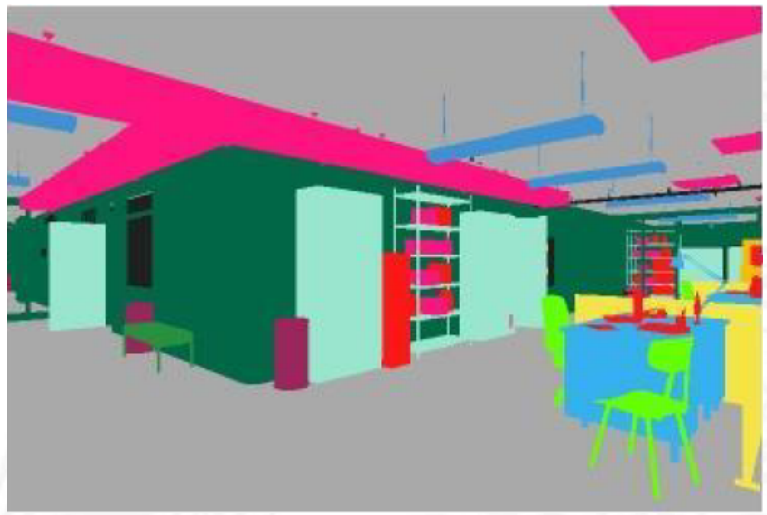
Source: [Rosinol et al., 2019a]

Semantics - Labels & Pointcloud

Pointcloud

Source: [Rosinol et al., 2019a]

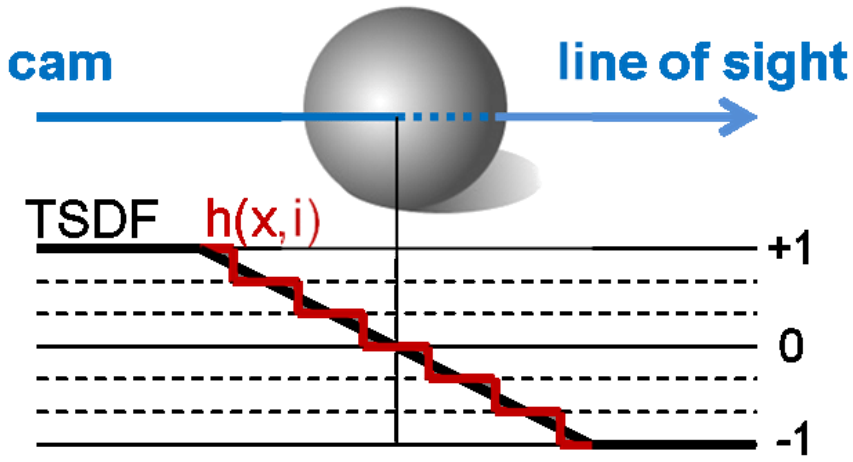
2D Labels



Source: [Rosinol et al., 2019a]

Semantics - TSDF

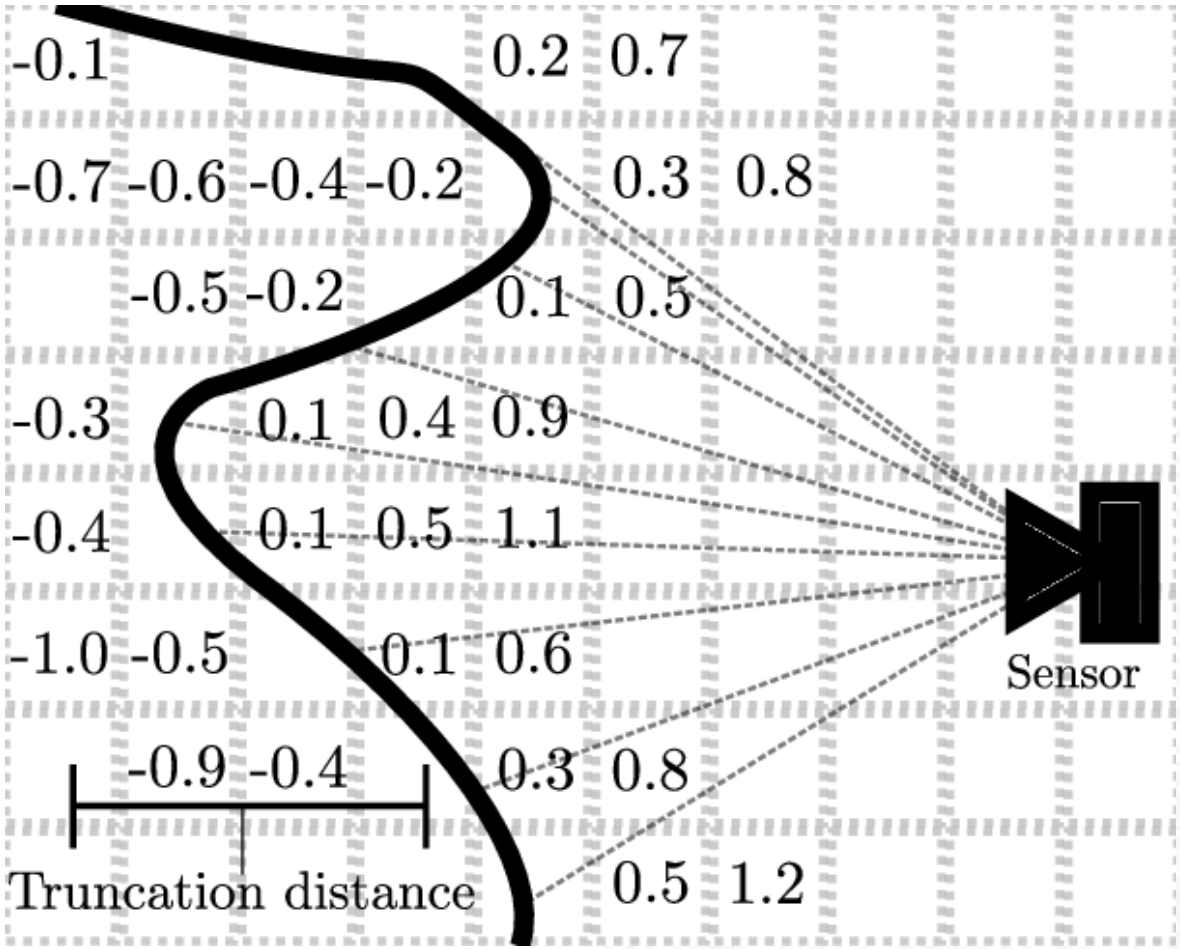
-0.9	-0.3	0.0	0.2	1	1	1	1	1
-1	-0.9	-0.2	0.0	0.2	1	1	1	1
-1	-0.9	-0.3	0.0	0.1	0.9	1	1	1
-1	-0.8	-0.3	0.0	0.2	0.8	1	1	1
-1	-0.9	-0.4	-0.1	0.1	0.8	0.9	1	1
-1	-0.7	-0.3	0.0	0.3	0.6	1	1	1
-1	-0.7	-0.4	0.0	0.2	0.7	0.8	1	1
-0.9	-0.7	-0.2	0.0	0.2	0.8	0.9	1	1
-0.1	0.0	0.0	0.1	0.3	1	1	1	1
0.5	0.3	0.2	0.4	0.8	1	1	1	1



Source: [Wendel and Wendel, 2016]

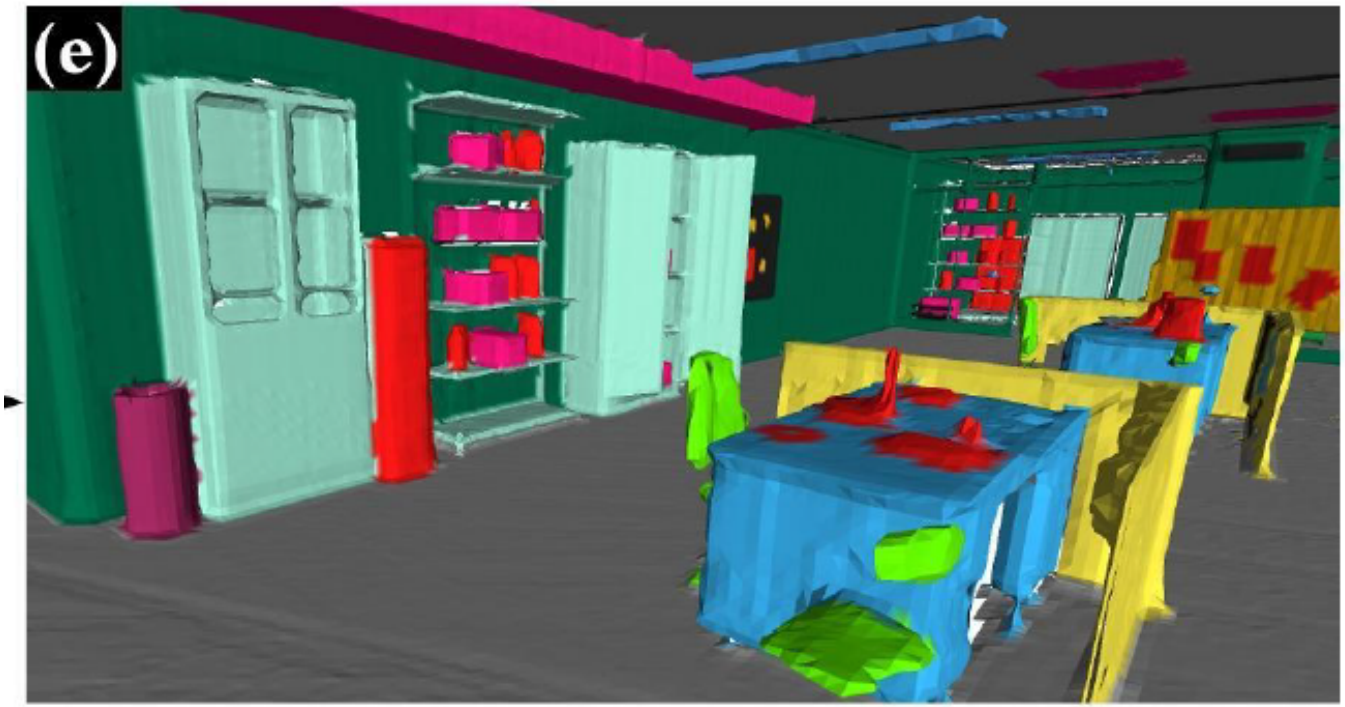
Source: https://community.arm.com/developer/tools-software/graphics/b/blog/posts/the-rise-of-depth-on-mobile?CommentId=6a22b56d-81d7-4be5-953e-7f1e3bdb2891?CommentSortBy=CreatedDateCommentSortOrder=Descending

Semantics - Raycasting



Source: [Whelan et al., 2015]

Semantics - Global Mesh



Source: [Rosinol et al., 2019a]

Conclusion

- One remarkable solution for a combination of highly complex problems
- Comes with a dense documentation in form of papers, videos and lectures
- **But,** it is not a Silver Bullet, there are other comparable solutions

References

[Amer et al., 2012] Amer, S. I., Eskander, M. N., and Zaki, A. M. (2012).
Positioning And Motion Control For Mobile Robot.
2(11):498–504.

[Castaldo et al., 2014] Castaldo, G., Angrisano, A., Gaglione, S., and Troisi, S. (2014).
P-RANSAC: An integrity monitoring approach for GNSS signal degraded scenario.
International Journal of Navigation and Observation, 2014(September).

References (cont.)

[Díaz Barros et al., 2019] Díaz Barros, J. M., Mirbach, B., Garcia, F., Varanasi, K., and Stricker, D. (2019).
Real-time head pose estimation by tracking and detection of keypoints and facial landmarks.
In *Communications in Computer and Information Science*, volume 997, pages 326–349.

[Forster et al., 2015] Forster, C., Carlone, L., Dellaert, F., and Scaramuzza, D. (2015).
On-Manifold Preintegration Theory for Fast and Accurate Visual-Inertial Navigation.
IEEE Transactions on Robotics.

References (cont.)

[Kaess et al., 2010] Kaess, M., Ila, V., Roberts, R., and Dellaert, F. (2010).
The Bayes tree: An algorithmic foundation for probabilistic robot mapping.
Springer Tracts in Advanced Robotics, 68(STAR):157–173.

[Kaess et al., 2012] Kaess, M., Johannsson, H., Roberts, R., Ila, V., Leonard, J. J., and Dellaert, F. (2012).
ISAM2: Incremental smoothing and mapping using the Bayes tree.
International Journal of Robotics Research, 31(2):216–235.

References (cont.)

[Oleynikova et al., 2017] Oleynikova, H., Taylor, Z., Fehr, M., Siegwart, R., and Nieto, J. (2017).
Voxblox: Incremental 3D Euclidean Signed Distance Fields for on-board MAV planning.
IEEE International Conference on Intelligent Robots and Systems, 2017-Septe:1366–1373.

[Rosinol et al., 2019a] Rosinol, A., Abate, M., Chang, Y., and Carlone, L. (2019a).
Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping.

References (cont.)

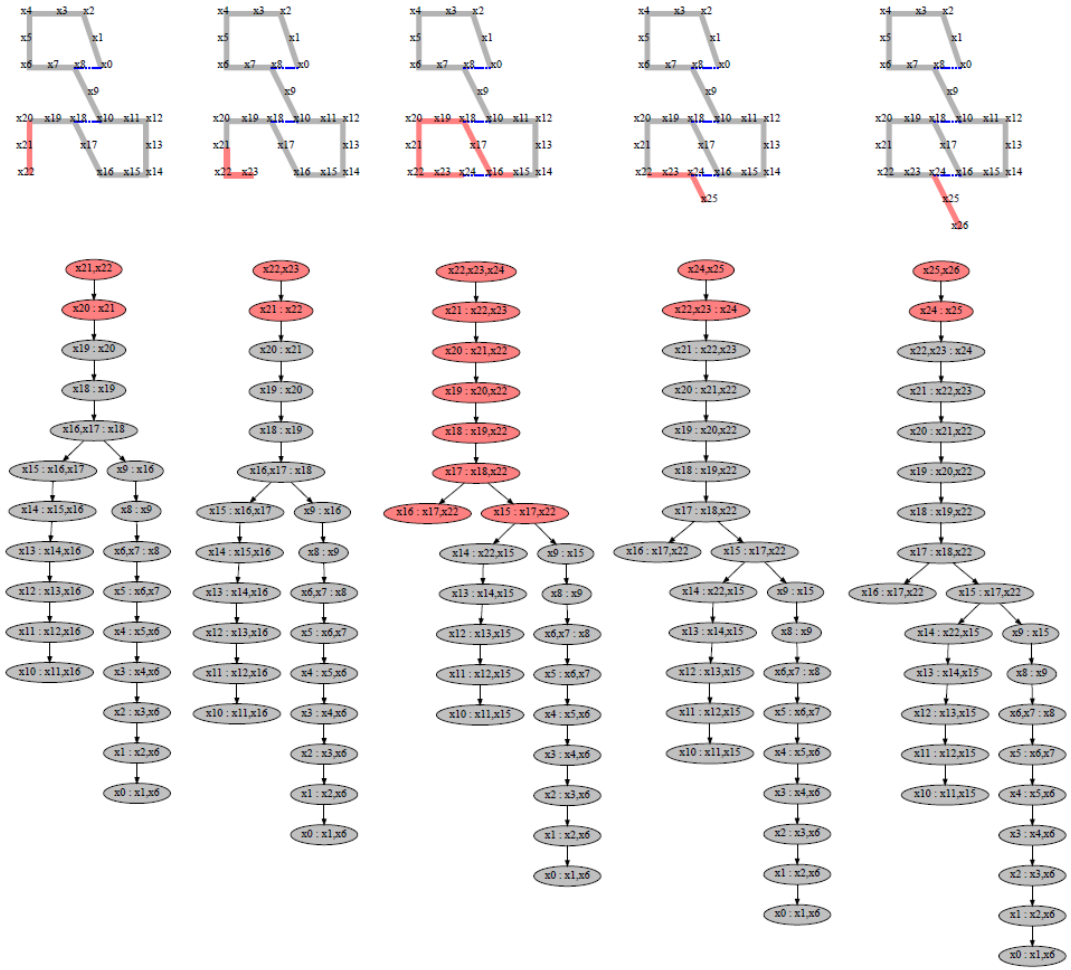
[Rosinol et al., 2019b] Rosinol, A., Sattler, T., Pollefeys, M., and Carlone, L. (2019b).
Incremental visual-inertial 3d mesh generation with structural regularities.
Proceedings - IEEE International Conference on Robotics and Automation, 2019-May(i):8220–8226.

[Wendel and Wendel, 2016] Wendel, A. and Wendel, A. (2016).
Scalable Visual Navigation for Micro Aerial Vehicles using Geometric Prior Knowledge Graz University of Technology Institute for Computer Graphics and Vision Dissertation Scalable Visual Navigation for Micro Aerial Vehicles using Geometric Prior Knowledge.
(December).

References (cont.)

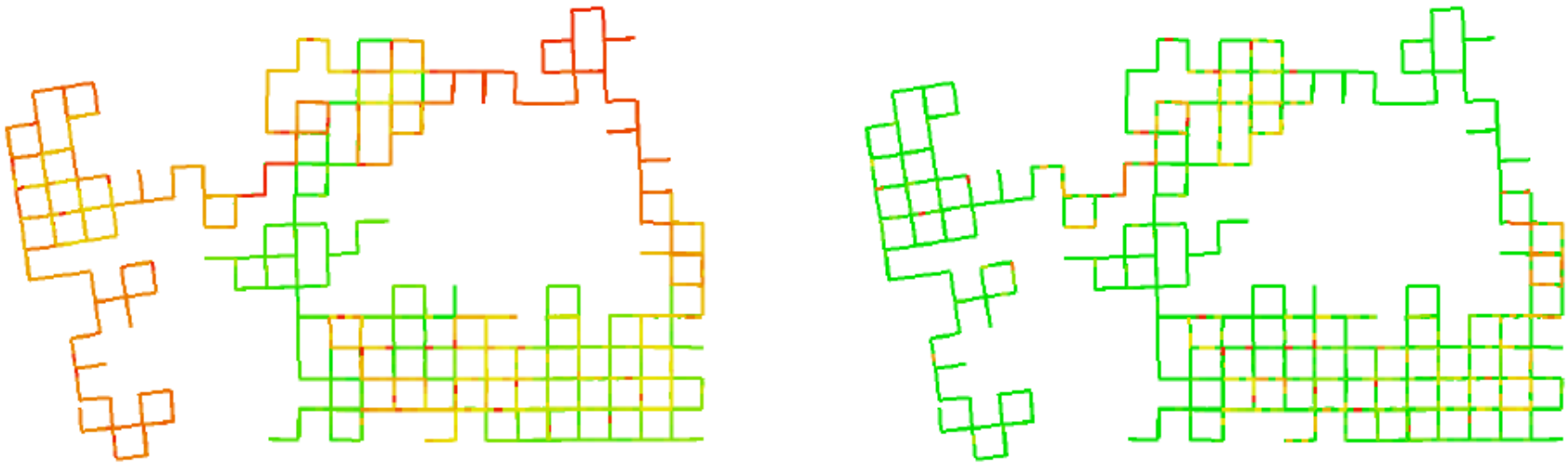
[Whelan et al., 2015] Whelan, T., Kaess, M., Johannsson, H., Fallon, M., Leonard, J. J., and McDonald, J. (2015).
Real-time large-scale dense RGB-D SLAM with volumetric fusion.
International Journal of Robotics Research.

Example: Bayes Tree



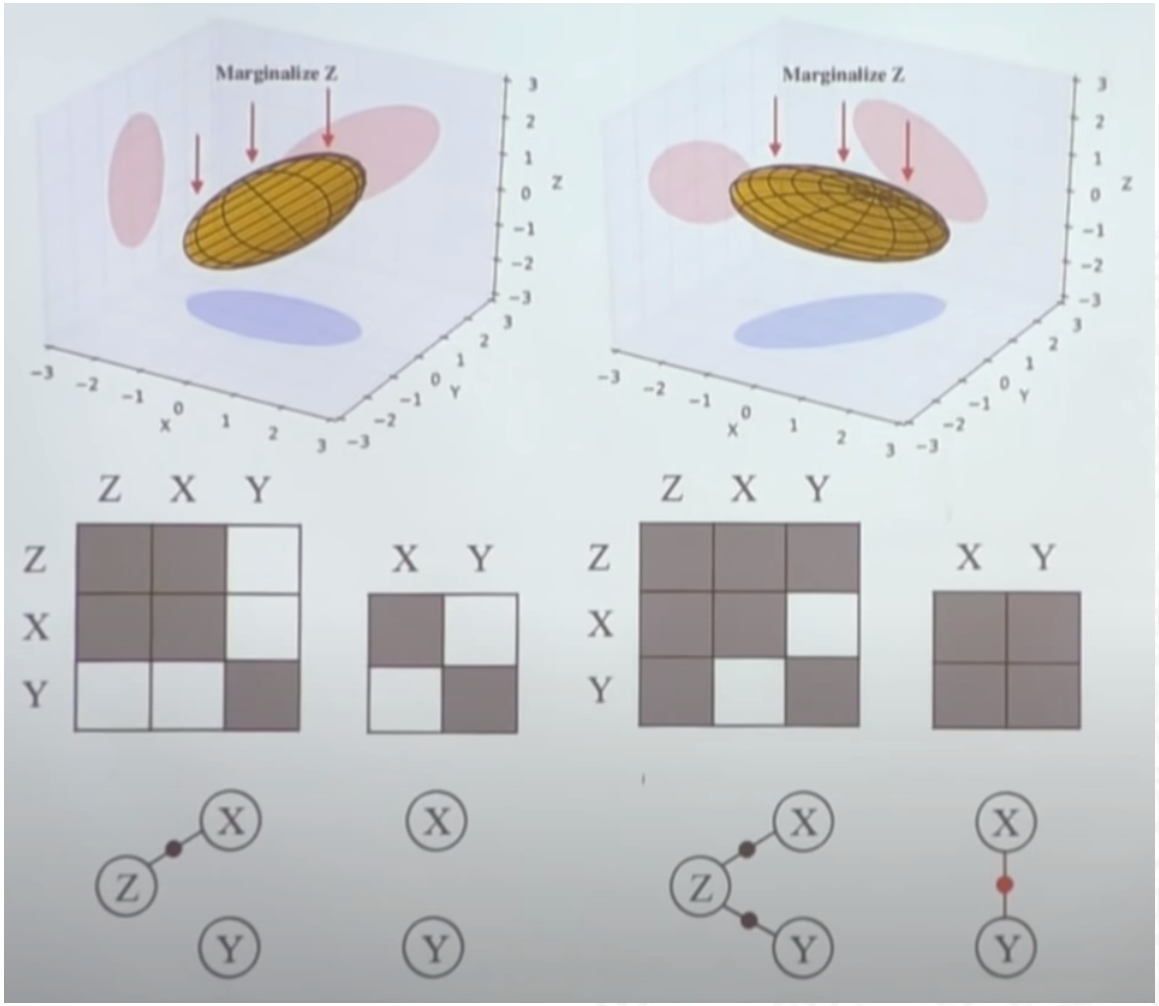
Source: [Kaess et al., 2012]

Example: Manhattan Problem



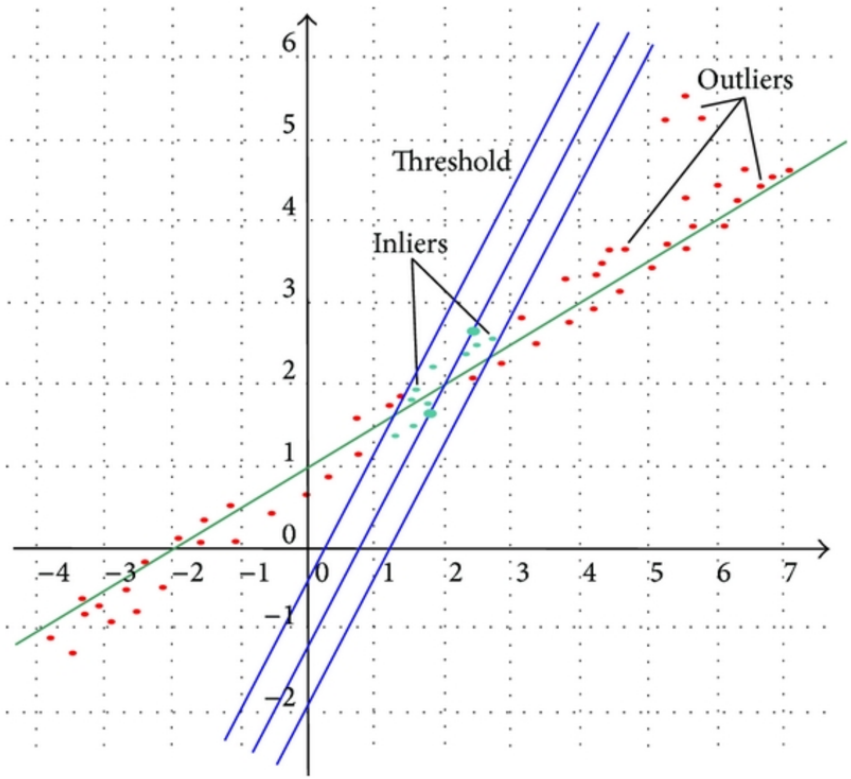
Source: [Kaess et al., 2012]

3D Marginalization



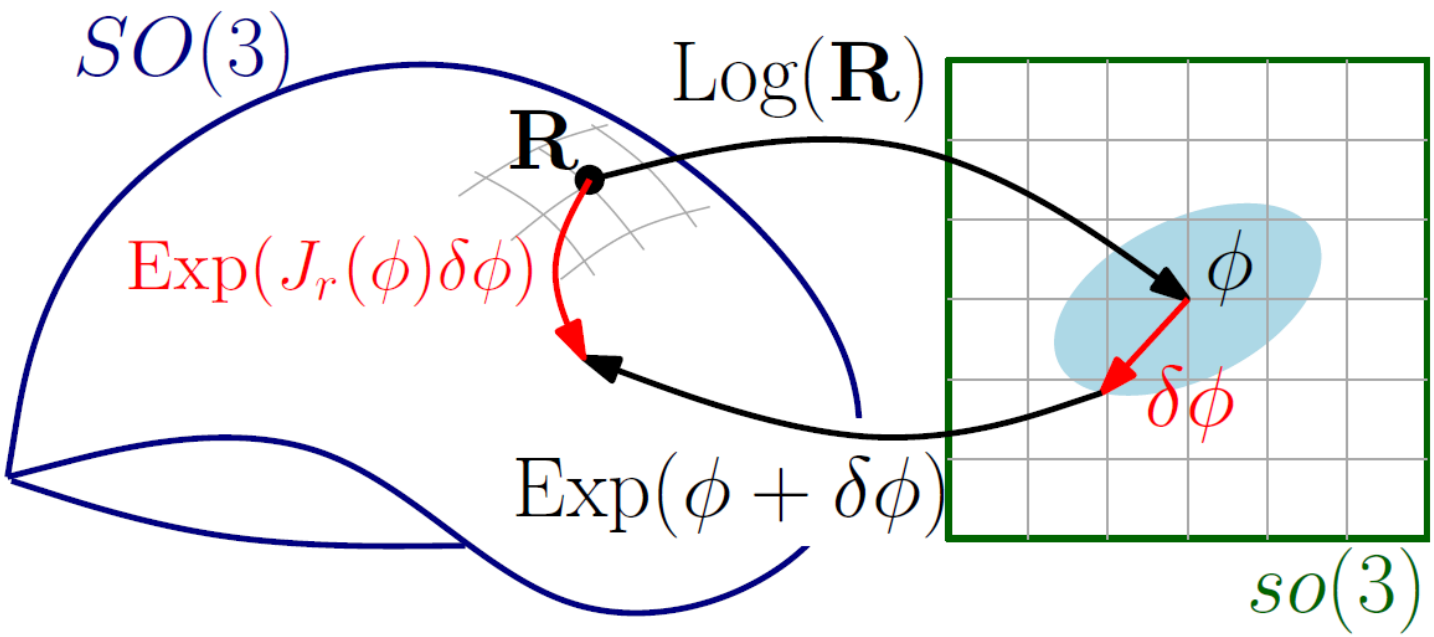
Source: https://www.youtube.com/watch?v=Q313pTMAdcM

RANSAC Algorithm



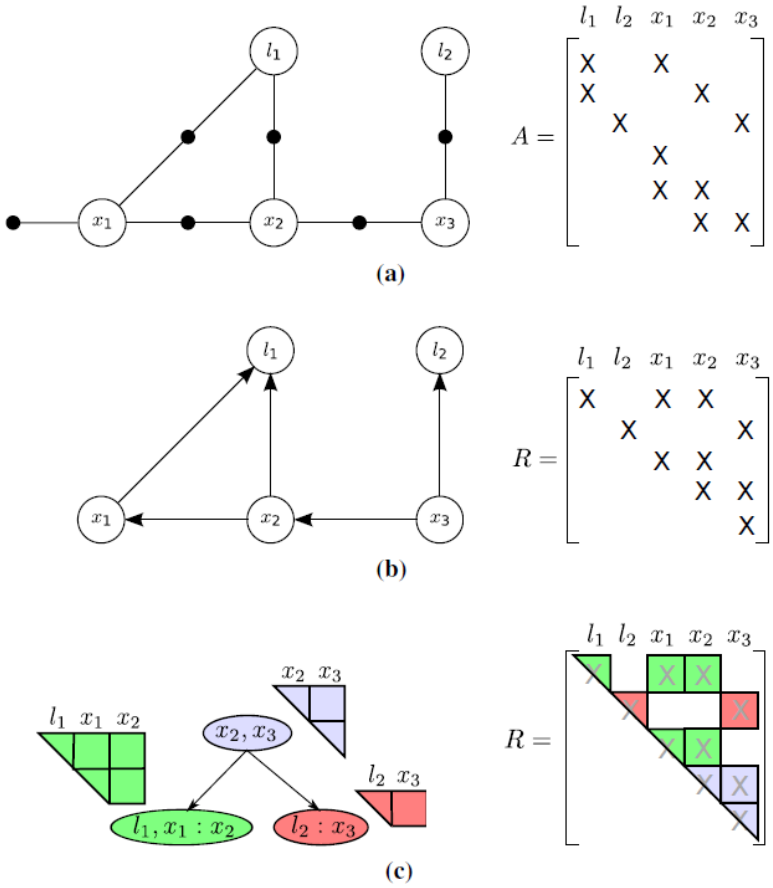
Source: [Castaldo et al., 2014]

IMU Preintegration - Tangential Plane



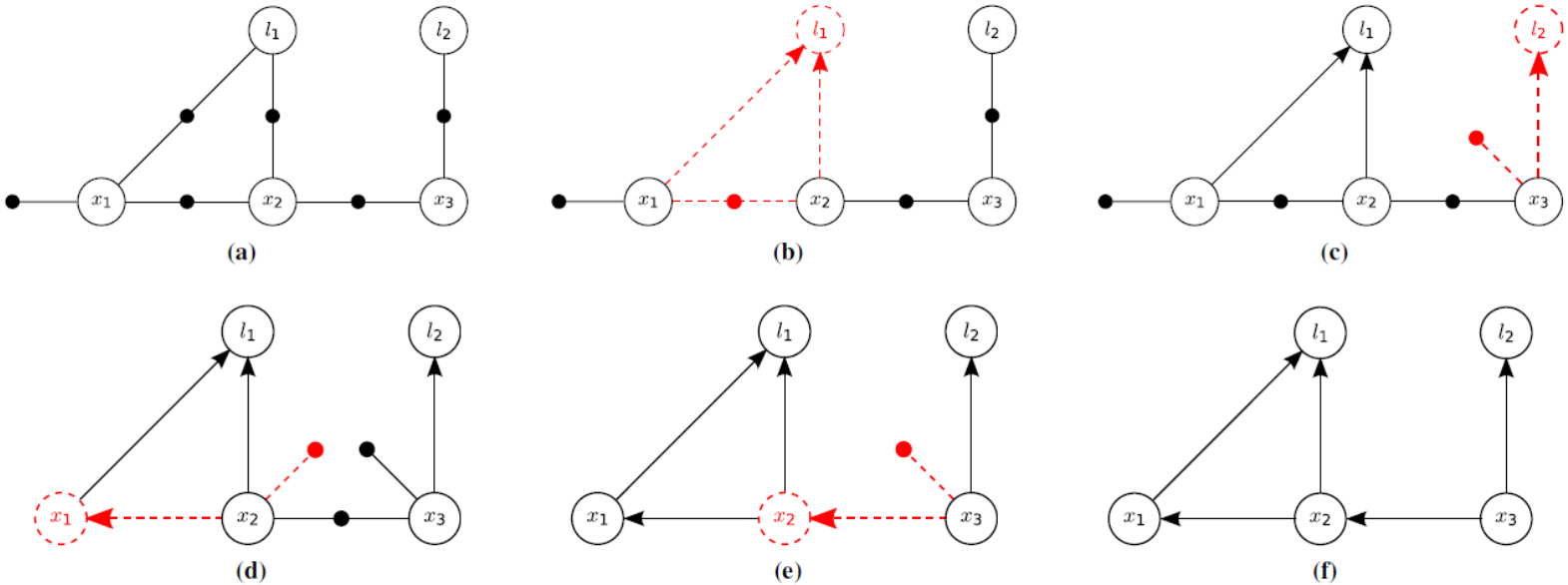
Source: [Forster et al., 2015]

Graphs and Trees in Matrix Form



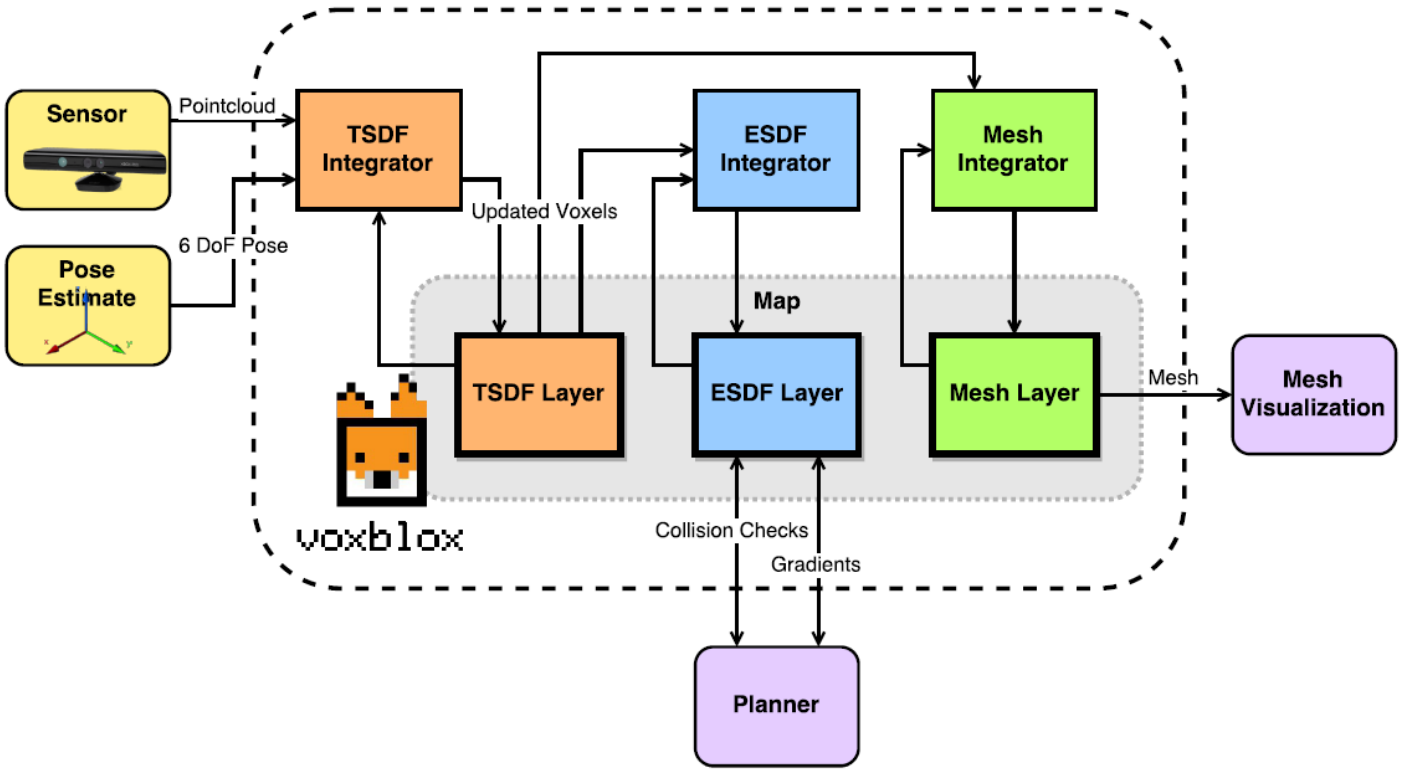
Source: [Kaess et al., 2012]

Example: Variable Elimination



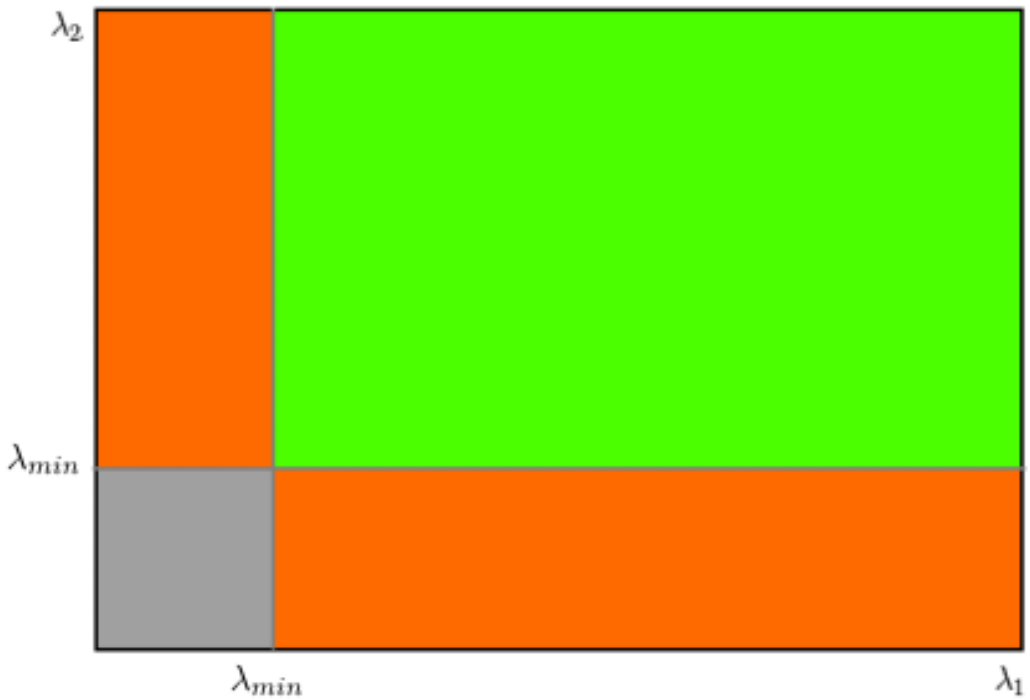
Source: [Kaess et al., 2012]

Voxblox



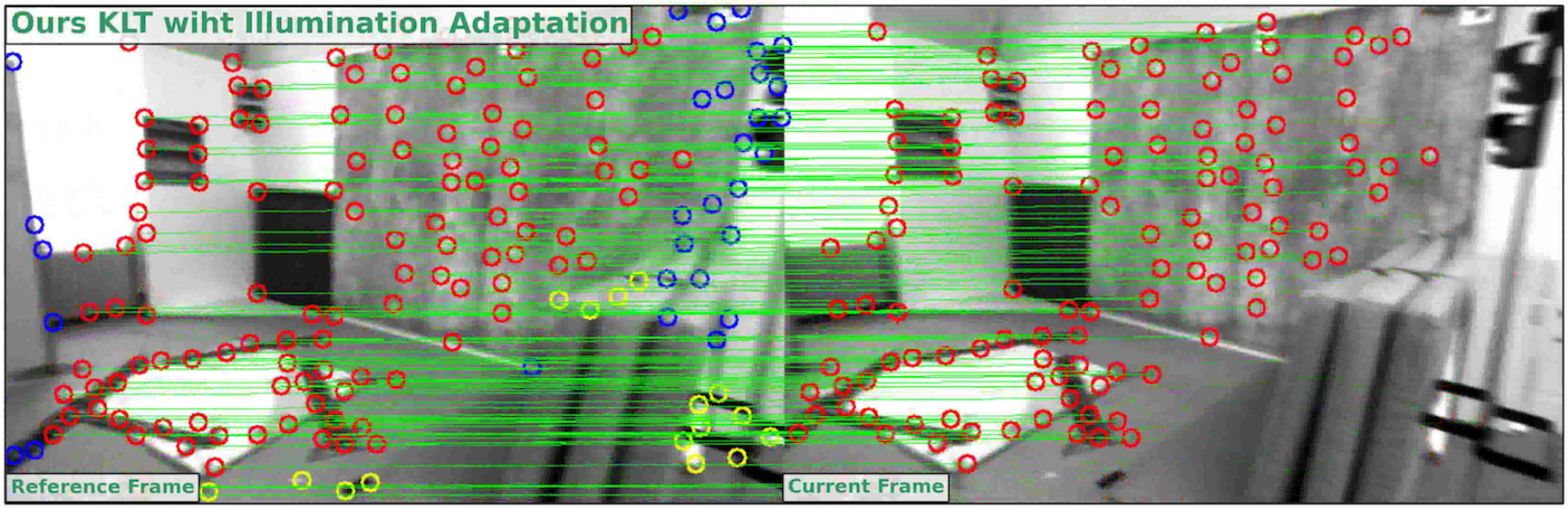
Source: [Oleynikova et al., 2017]

Shi-Tomasi Corner Detection



Source: https://docs.opencv.org/3.4/d4/d8c/tutorial_py_shi_tomasi.html

Lukas-Kanade Tracker



Source: http://www.luohanjie.com/2018-04-13/klt-with-illumination-adaptation.html