Online large-scale SLAM with stereo visual-inertial sensors

Dominik Schlegel

Supervisor: Prof. Dr. Giorgio Grisetti

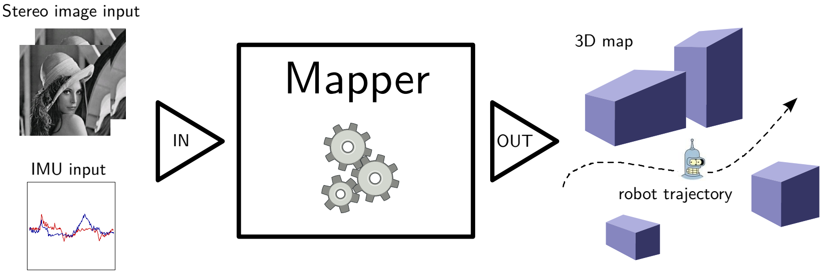
November 12, 2015

ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

SAPIENZA
UNIVERSITÀ DI ROMA

Aim of this project

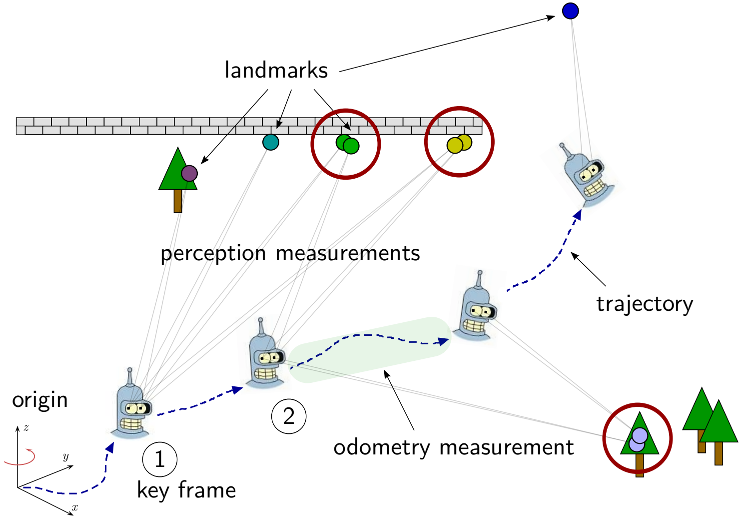
Map the world using a stereo camera and an IMU:



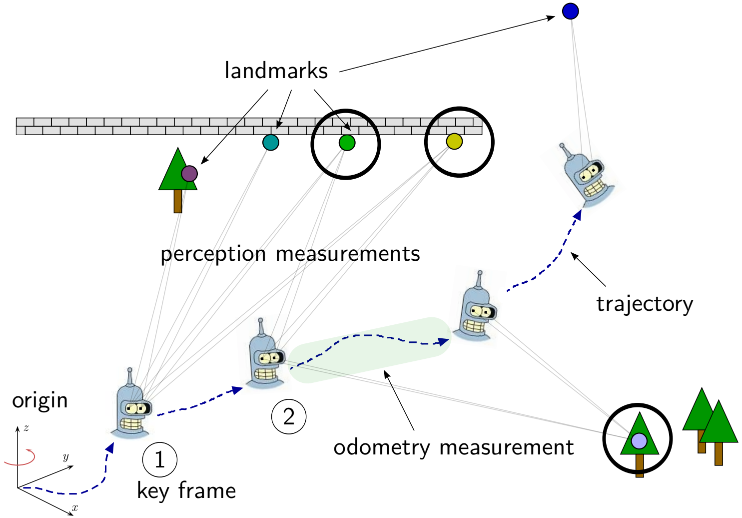
Required capabilities:

- Tracking (Localization)
- Mapping (Solving SLAM problem)

Graph-based SLAM: The problem



Graph-based SLAM: The solution with g2o[1]



[1] **Kuemmerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W.: g2o: A general framework for graph optimization (ICRA 2011).**

State of the art

Google (street) mapping

Car

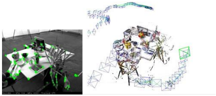
Trekker

Trolley

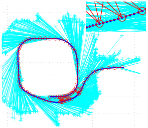
Snow mobile

Trike

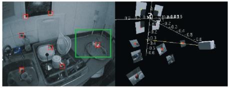
ORB-SLAM



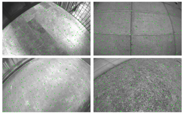
FrameSLAM



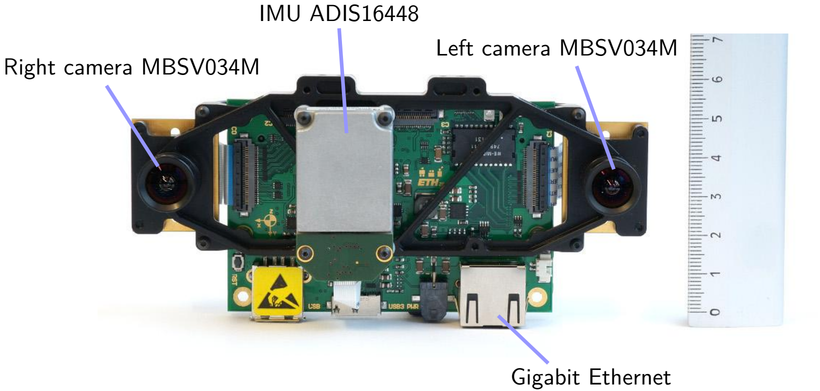
MonoSLAM



SVO

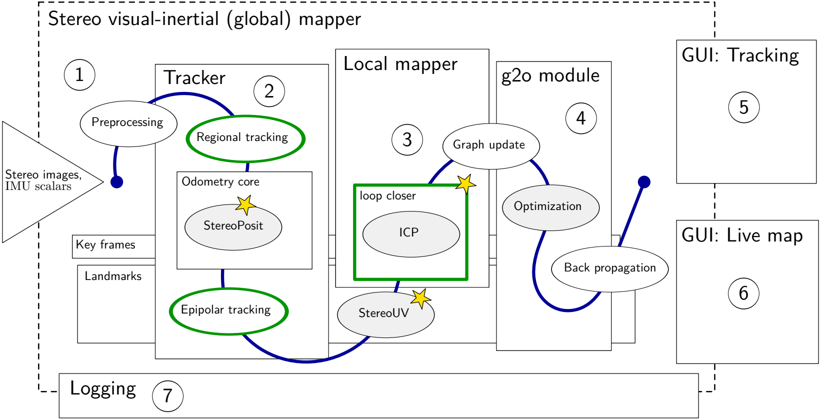


Sensor setup: the VI-Sensor[2]



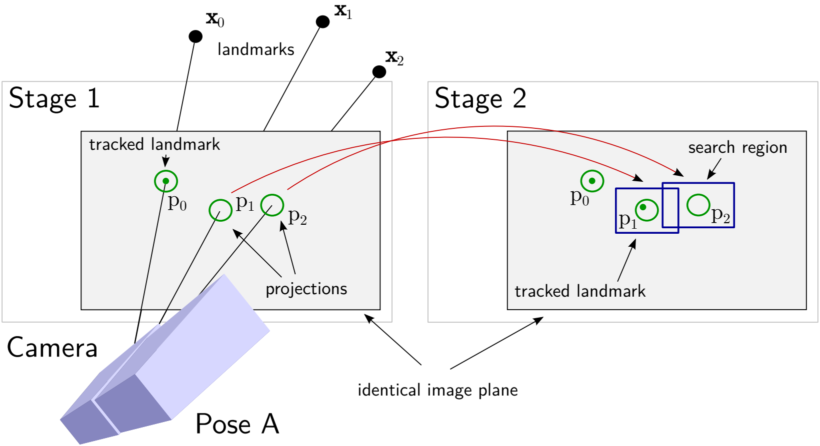
[2] **Nikolic**, **J.**, **Rehder**, **J.**, **Burri**, **M.**, **Gohl**, **P.**, **Leutenegger**, **S.**, **Furgale**, **P.T.**, **Siegwart**, **R.**: **A synchronized visual-inertial sensor system with FPGA pre-processing for accurate real-time SLAM (ICRA 2014)**.

Our system



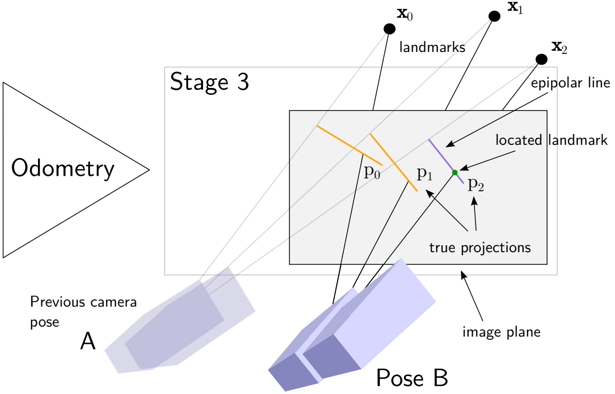
Contributions: StereoPosit, StereoUV, Loop closure detection

Tracking: Regional tracking[3]



[3] **Calonder**, **M.**, **Lepetit**, **V.**, **Strecha**, **C.**, **Fua**, **P.**: **Brief: Binary robust independent elementary features. Computer Vision-ECCV 2010**.

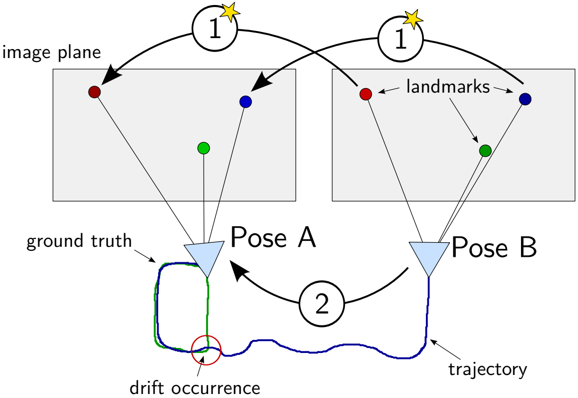
Tracking: Epipolar tracking[4]



In action: Aula Magna

[4] **Richard**, **Hartley; Andrew**,
Z.: Multiple View Geometry in Computer Vision. Cambridge University Press (March 2004).

Loop closing: Our approach



- Step 1: Find matching landmarks (Contribution)
- Step 2: Compute spatial relation: ICP[5]

Aula Magna

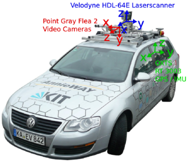
[5] **Besl, P.J., McKay, N.D.: Method for registration of 3-D shapes. Proc. SPIE 1611 (1992)**.

Data acquisition[6]

Hand-held

Bike

KITTI Annieway

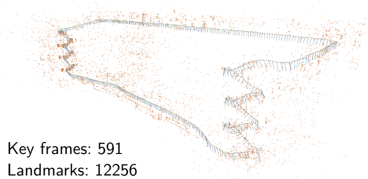


[6] **Geiger**, **A.**, **Lenz**, **P.**, **Urtasun**,
R.: Are we ready for autonomous driving? the KITTI vision benchmark suite (CVPR 2012).

Results: Hand-held

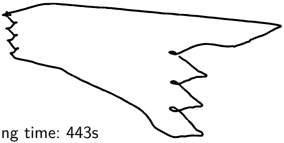
Stairs DIAG

Full graph



Key frames: 591
Landmarks: 12256

Trajectory



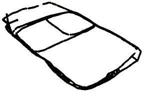
Recording time: 443s
Trajectory length: 394m

Aula Magna

Full graph

Key frames: 213
Landmarks: 4049

Trajectory



Recording time: 192s
Trajectory length: 126m

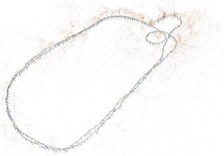
Results: Bike

Parco San Lorenzo

Key frames: 657
Landmarks: 17126

Recording time: 497s
Trajectory length: 381m

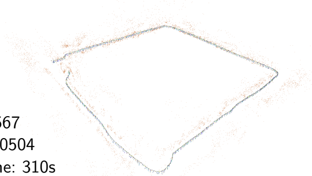
Full graph | Trajectory (Google maps overlay)

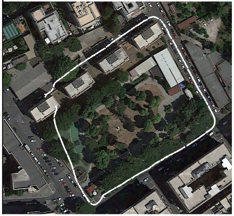


Block San Lorenzo

Key frames: 567
Landmarks: 10504
Recording time: 310s
Trajectory length: 354m

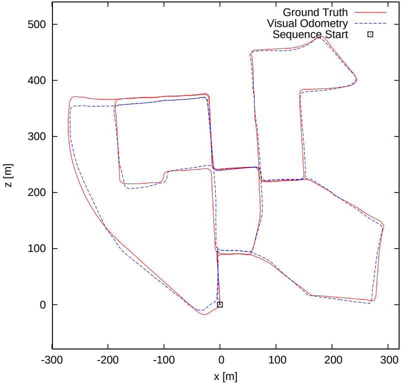
Full graph | Trajectory (Google maps overlay)



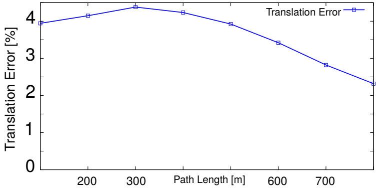


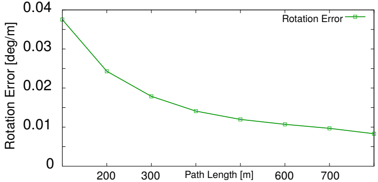
Results: KITTI odometry benchmark (without IMU)

Trajectory overlay:



Errors (KITTI standard):





Total trajectory length: 3722m

Conclusions and remarks

Final system:

- Outstanding robustness
- High accuracy (Map overlays, 1-5% KITTI error)
- Strong, independent odometry (KITTI)
- Modularity (Multithreading)

Future work:

- Filter integration[7] (IMU)
- Improved feature handling[8]
- Improved graph optimization[9]

---

[7] **Forster**, **C.**, **Carlone**, **L.**, **Dellaert**, **F.**, **Scaramuzza**, **D.: IMU Preintegration on Manifold for Efficient Visual-Inertial Maximum-a-Posteriori Estimation. (RSS 2015)**.

[8] **Rublee**, **E.**, **Rabaud**, **V.**, **Konolige**, **K.**, **Bradski**, **G.: ORB: an efficient alternative to SIFT or SURF. (ICCV 2011)**.

[9] **Grisetti**, **Giorgio**, **Rainer Kuemmerle**, **and Kai Ni. Robust optimization of factor graphs by using condensed measurements. (IROS 2012)**.

Online large-scale SLAM with stereo visual-inertial sensors

Dominik Schlegel
Supervisor: Prof. Dr. Giorgio Grisetti

November 12, 2015

Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Sapienza Università di Roma