SLAM using hand-held cameras only

José Neira
Universidad de Zaragoza

Joint work with:
César Cadena, Andrew Davison, Lina Paz,
Pedro Piníes, Ian Reid, Juan Tardós,
Brian Williams

ROBOTICS
UNIVERSIDAD DE ZARAGOZA

Zaragoza, where is that?

Santiago Ramón y Cajal

Francisco de Goya

Motivation (late 1098s)

SLAM seeks to answer this question:

Is it possible to use a vehicle, starting at an

- **unknown initial location**, in an
- **unknown environment**, to
- **incrementally**

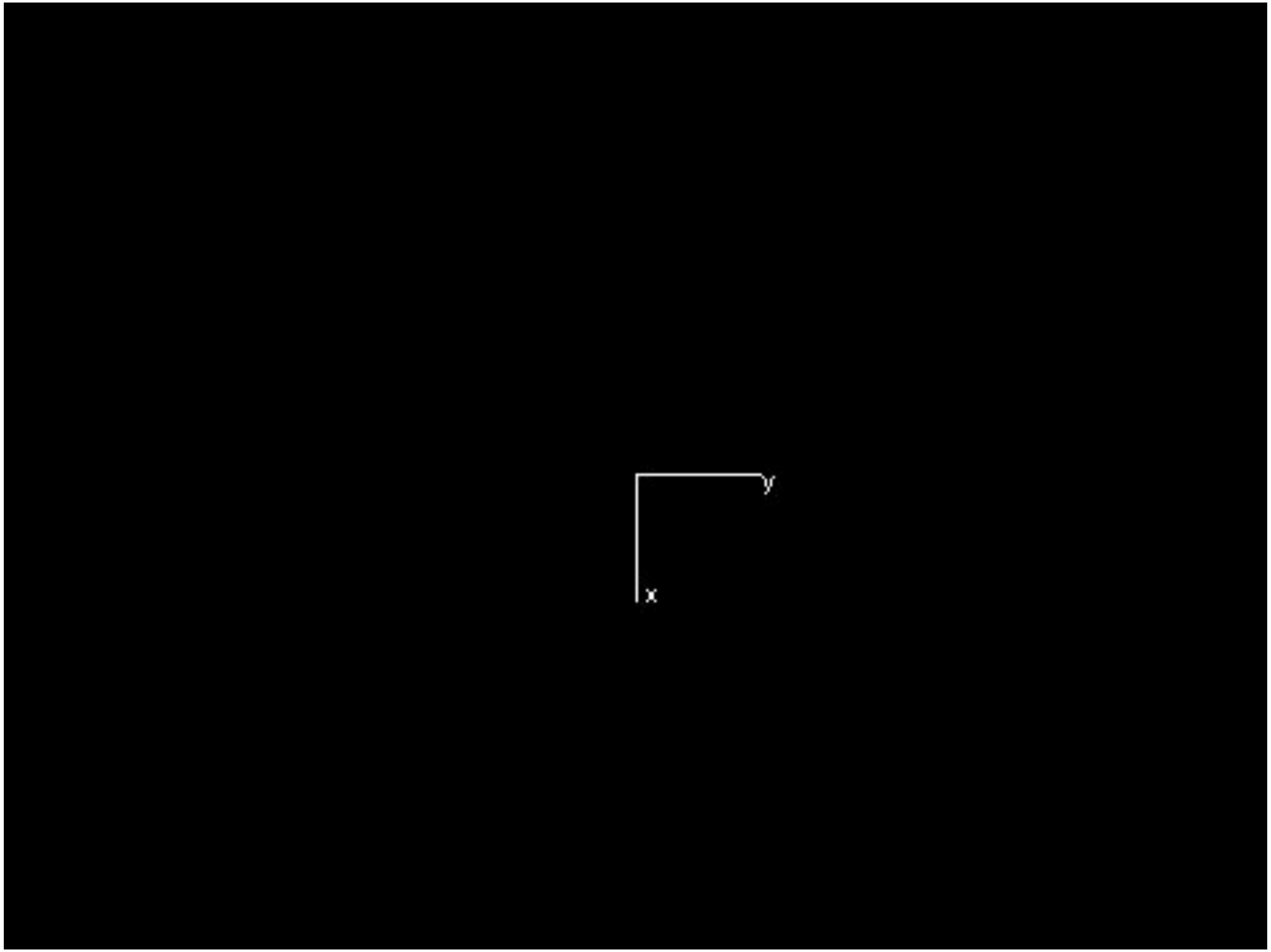
build a map of the environment,

- and **at the same time**

use the map to determine
the vehicle location?

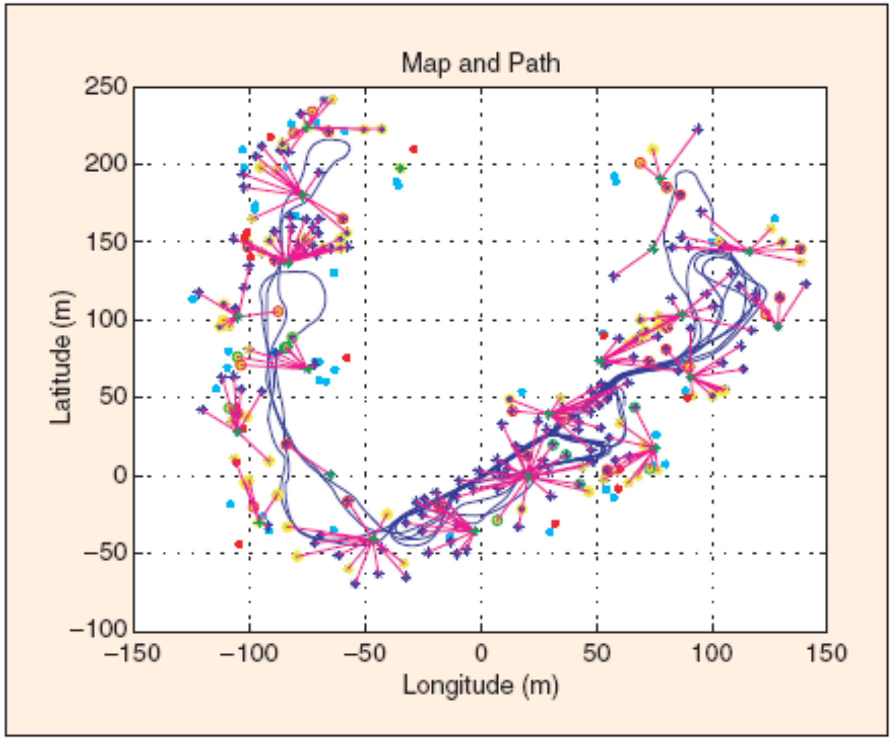
ROBOTICS
UNIVERSIDAD DE ZARAGOZA

Motivation



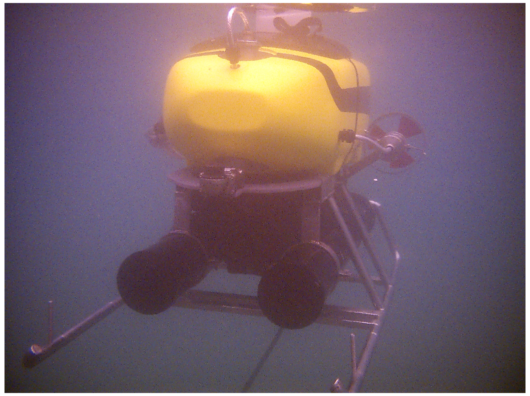
(video: Paul Newman)

Outdoor vehicles



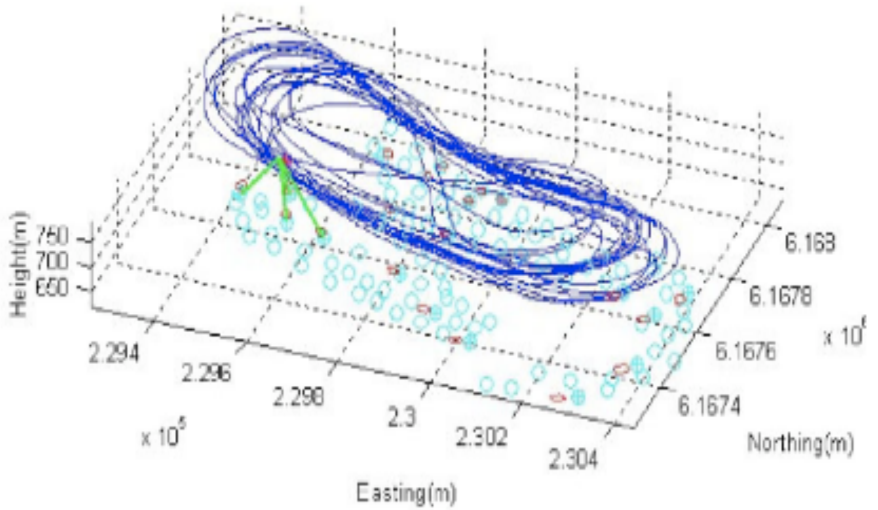
Victoria Park, Univ. Syndey

Underwater, Airborne



Garbi, Univ. Girona, Spain

Brumby, Univ. Syndey



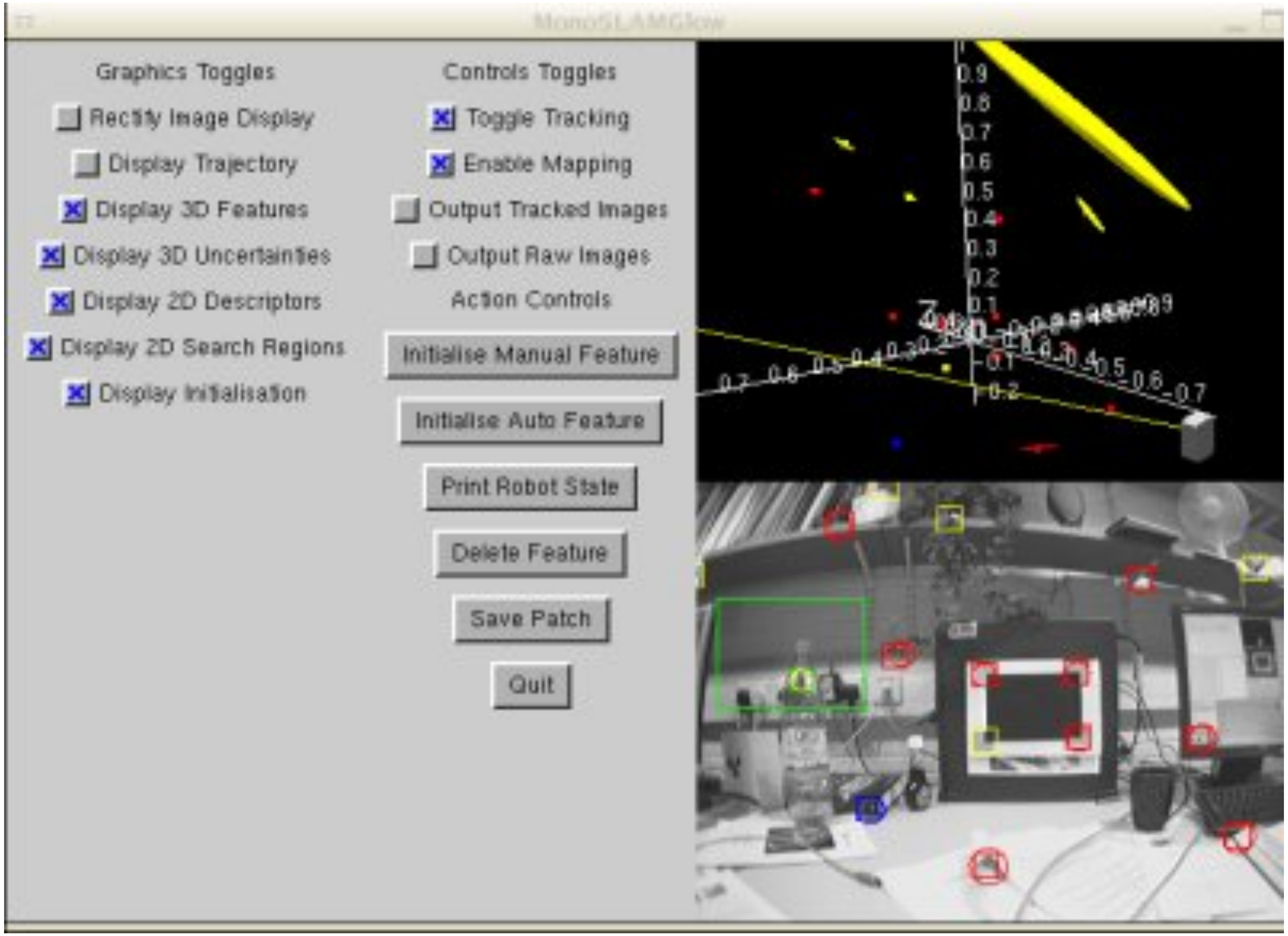
Fundamental issues

- **Scaling**: how large are the environments that we can map in real time?
- **Robustness**: can we deal with sensor error and cluttered environments?
- **Visual SLAM**: cameras are inexpensive, lightweight and provide enormous detail, can we do SLAM with cameras only?

Monocular SLAM

A Unibrain fire-I camera, a laptop and a firewire cable

Monoslam (A. Davison)



290 m.


© 2007 Infoterra Ltd & Bluesky
Google
Puntero 51°45'32.58" N 1°15'28.30" O
Secuencia 100%
Alt. ojo 116 m

The EKF SLAM algorithm

Algorithm 1 SLAM:

$\mathbf{x}_0^B = \mathbf{0}$; $\mathbf{P}_0^B = \mathbf{0}$ $\{$*Map initialization*$\}$

$[\mathbf{z}_0, \mathbf{R}_0] = \text{get\_measurements}$

$[\mathbf{x}_0^B, \mathbf{P}_0^B] = \text{add\_new\_features}(\mathbf{x}_0^B, \mathbf{P}_0^B, \mathbf{z}_0, \mathbf{R}_0)$

for $k = 1$ to steps **do**

$[\mathbf{x}_{R_k}^{R_{k-1}}, \mathbf{Q}_k] = \text{get\_odometry}$

$[\mathbf{x}_{k|k-1}^B, \mathbf{P}_{k|k-1}^B] = \text{EKF\_prediction}(\mathbf{x}_{k-1}^B, \mathbf{P}_{k-1}^B, \mathbf{x}_{R_k}^{R_{k-1}}, \mathbf{Q}_k)$

$[\mathbf{z}_k, \mathbf{R}_k] = \text{get\_measurements}$

$\mathcal{H}_k = \text{data\_association}(\mathbf{x}_{k|k-1}^B, \mathbf{P}_{k|k-1}^B, \mathbf{z}_k, \mathbf{R}_k)$

O(n²) $[\mathbf{x}_k^B, \mathbf{P}_k^B] = \text{EKF\_update}(\mathbf{x}_{k|k-1}^B, \mathbf{P}_{k|k-1}^B, \mathbf{z}_k, \mathbf{R}_k, \mathcal{H}_k)$

$[\mathbf{x}_k^B, \mathbf{P}_k^B] = \text{add\_new\_features}(\mathbf{x}_k^B, \mathbf{P}_k^B, \mathbf{z}_k, \mathbf{R}_k, \mathcal{H}_k)$

end for

3D features representation

3D points:

- - Cartesian coordinates

$$\mathbf{y}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \end{pmatrix}$$

Inverse depth points:

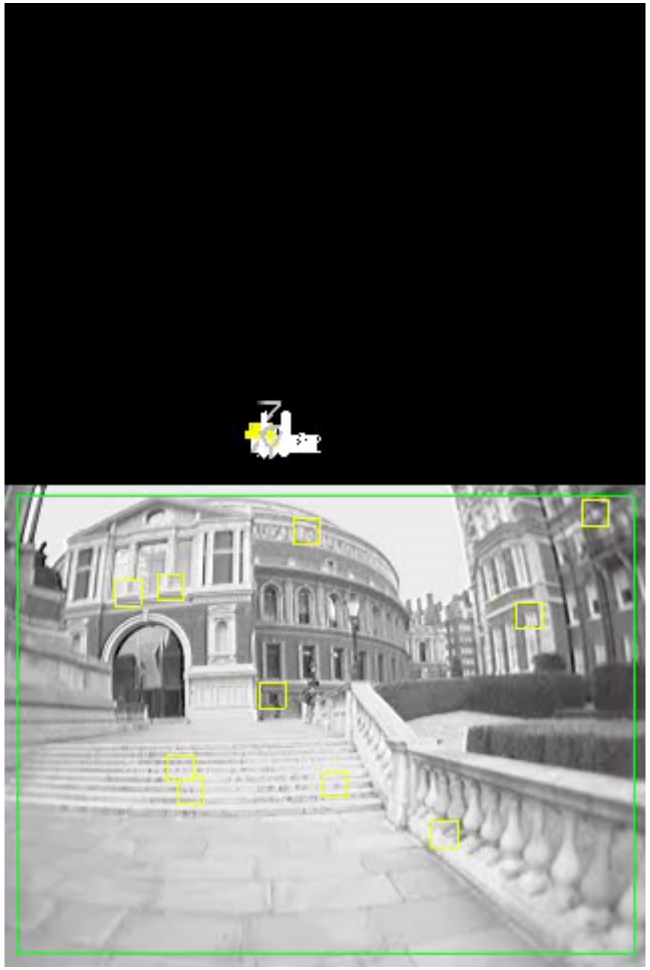
- Camera position the first time the feature was seen

- Azimuth
- Elevation
- Inverse depth

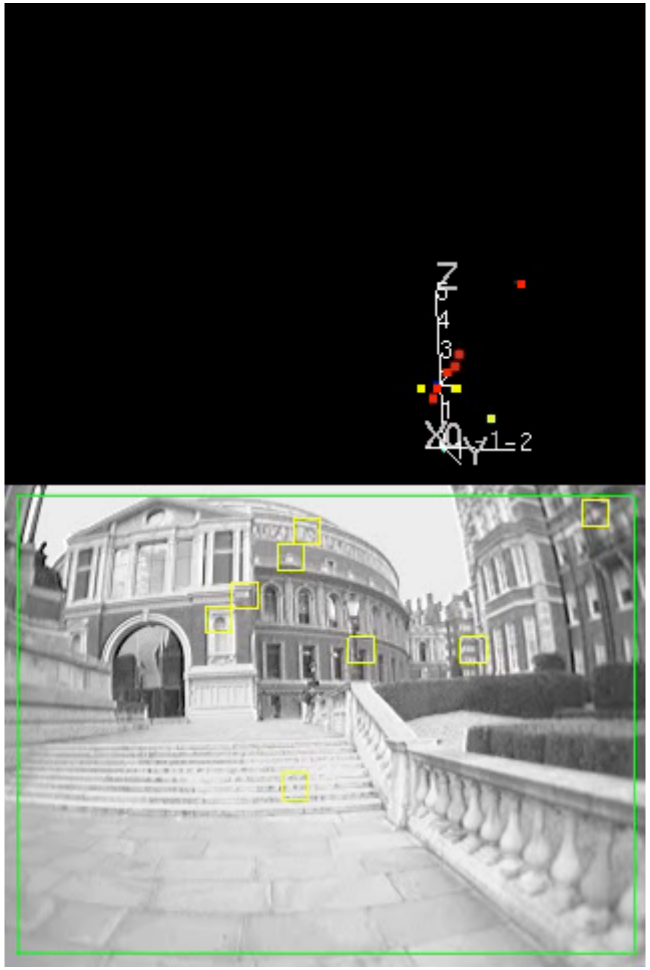
$$\mathbf{y}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \\ \theta_i \\ \phi_i \\ \rho_i \end{pmatrix}$$

J.M.M. Montiel, J. Civera, A.J. Davison: **Unified inverse depth parametrization for monocular SLAM.** IEEE Transactions on Robotics, October 2008.

Robustness: data association

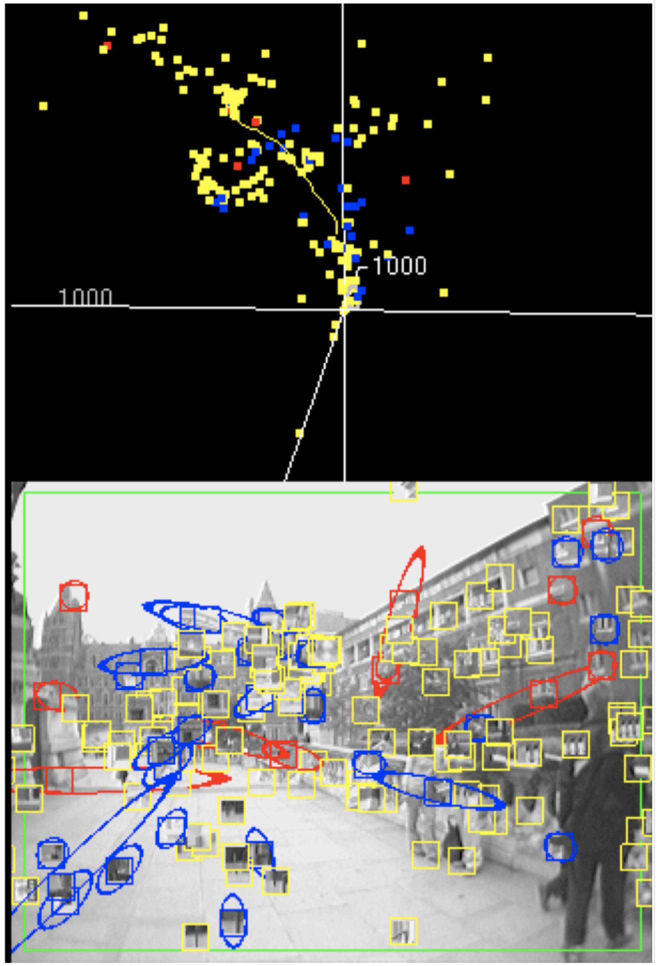


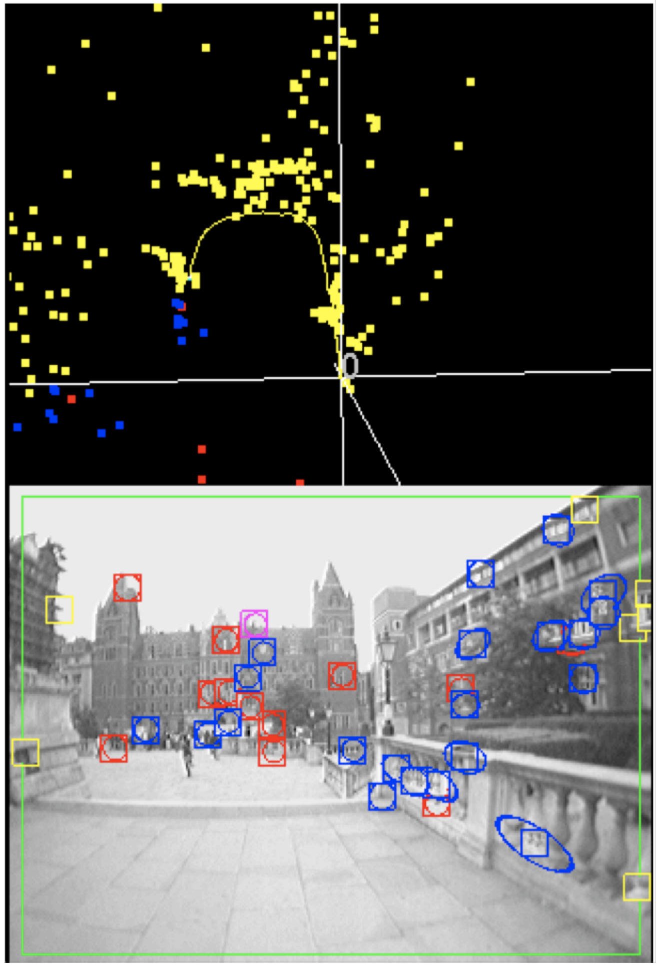
Individual tracks



Jointly compatible tracks

Nearest neighbor .vs. Joint Compatibility



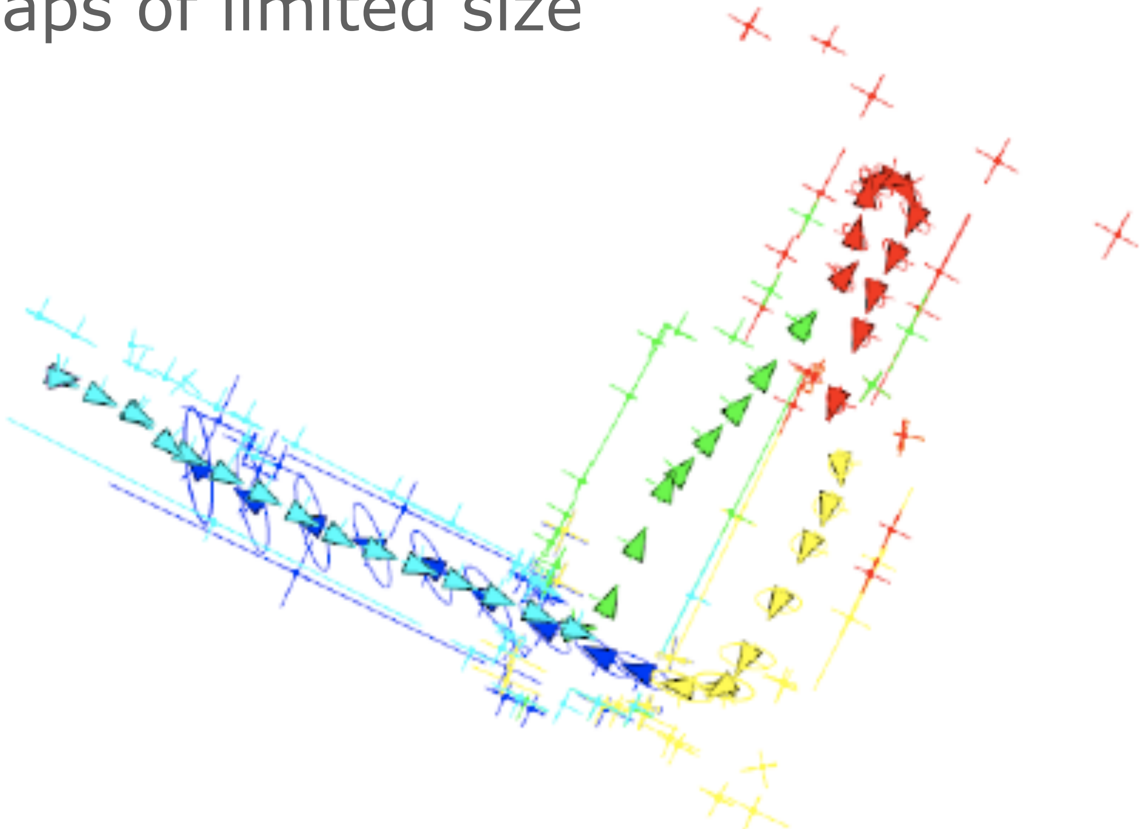


Individual Tracks

Jointly Compatible Tracks (cost: 2ms)

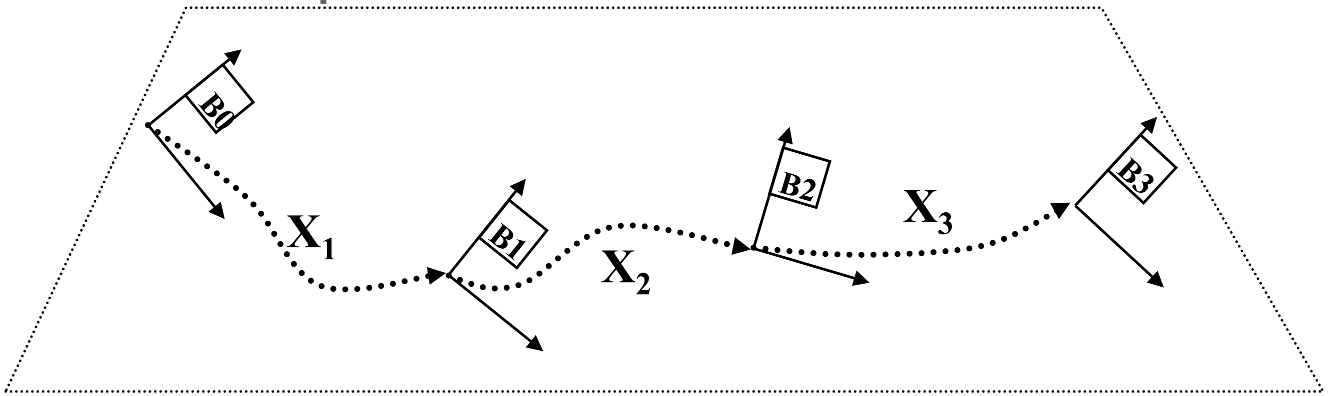
Scalable EKF SLAM:

- Local maps of limited size

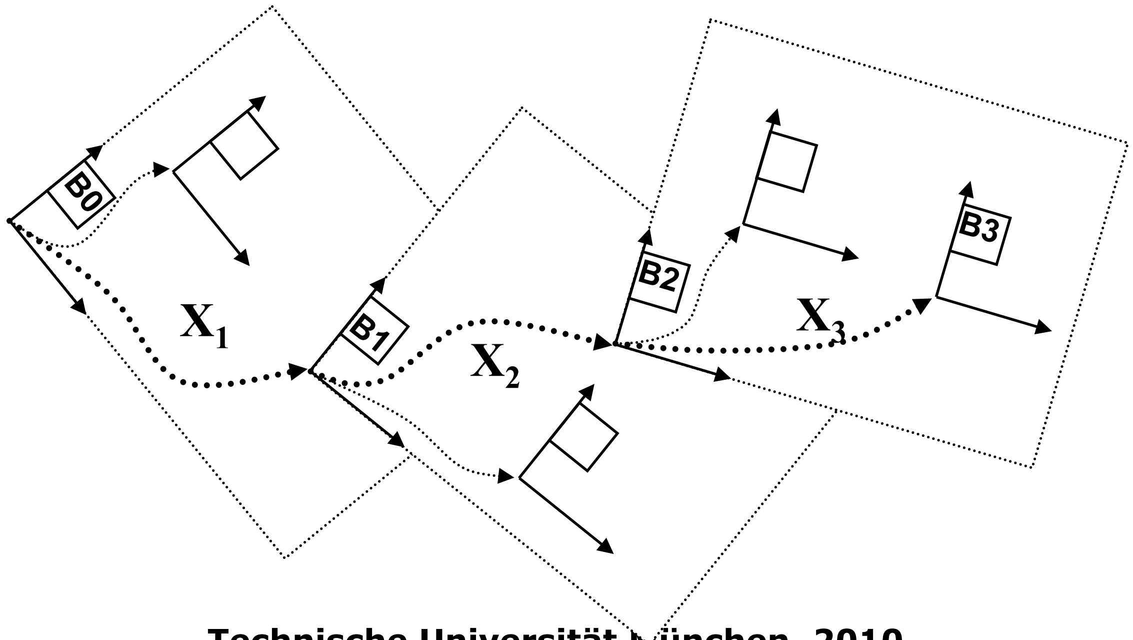


Hierarchical SLAM

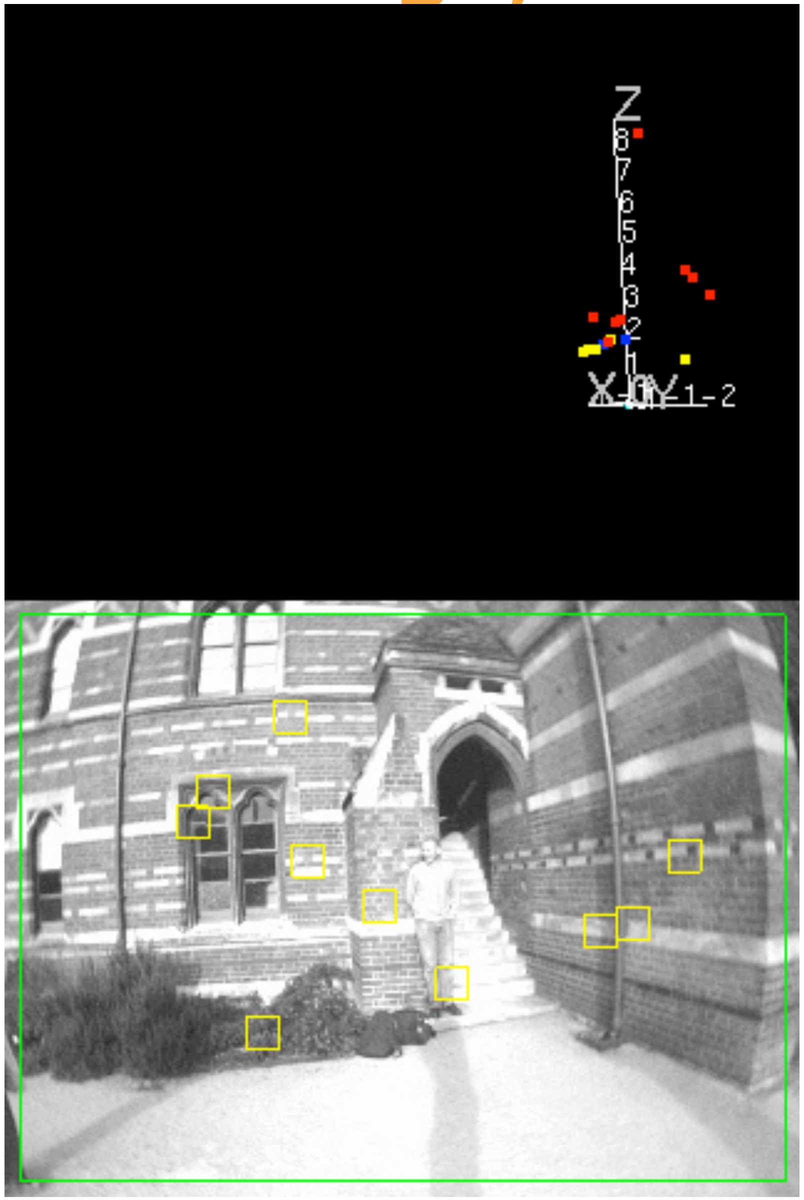
- Global level: adjacency graph and relative stochastic map



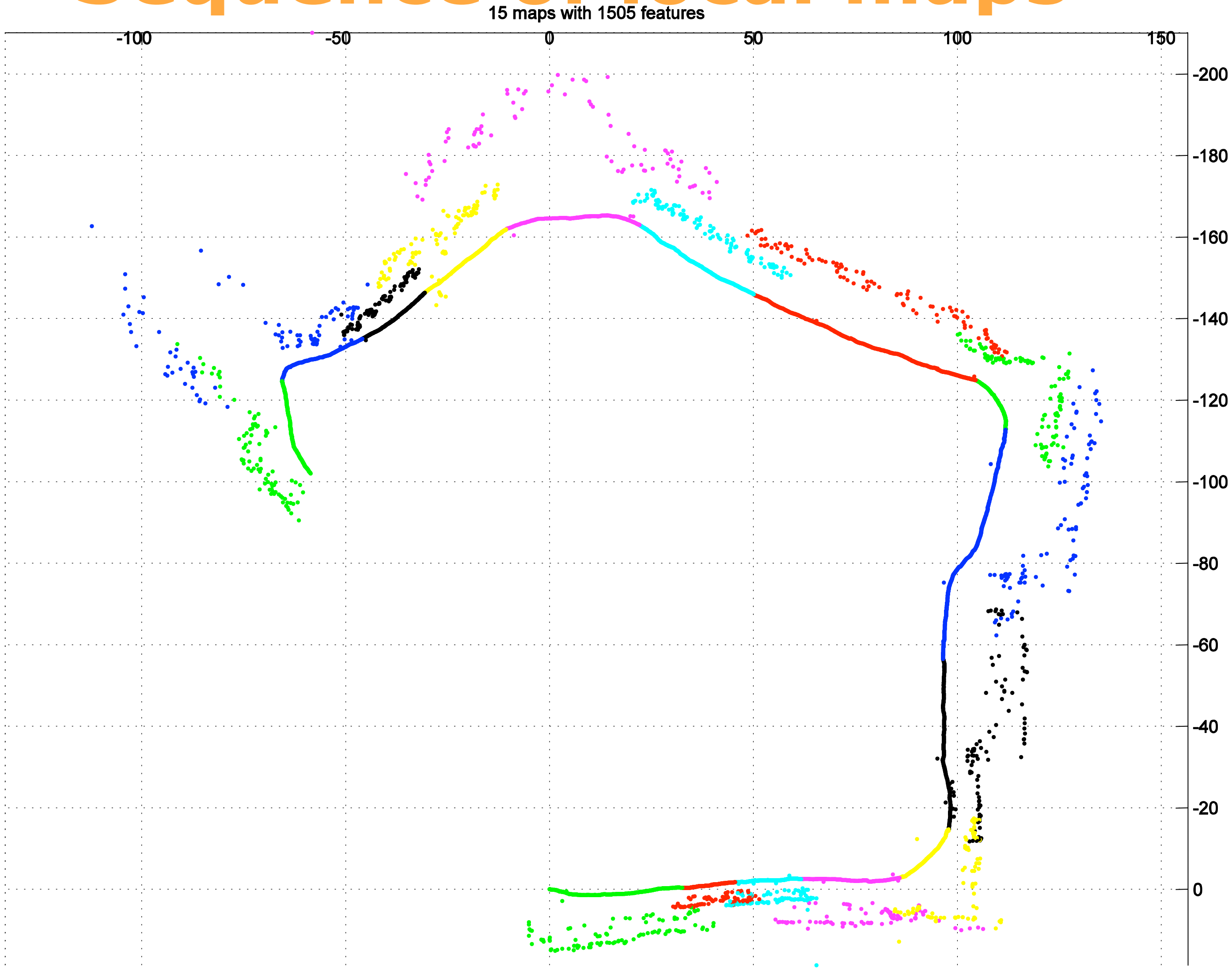
- Local level: statistically independent local maps



Keble College, Oxford

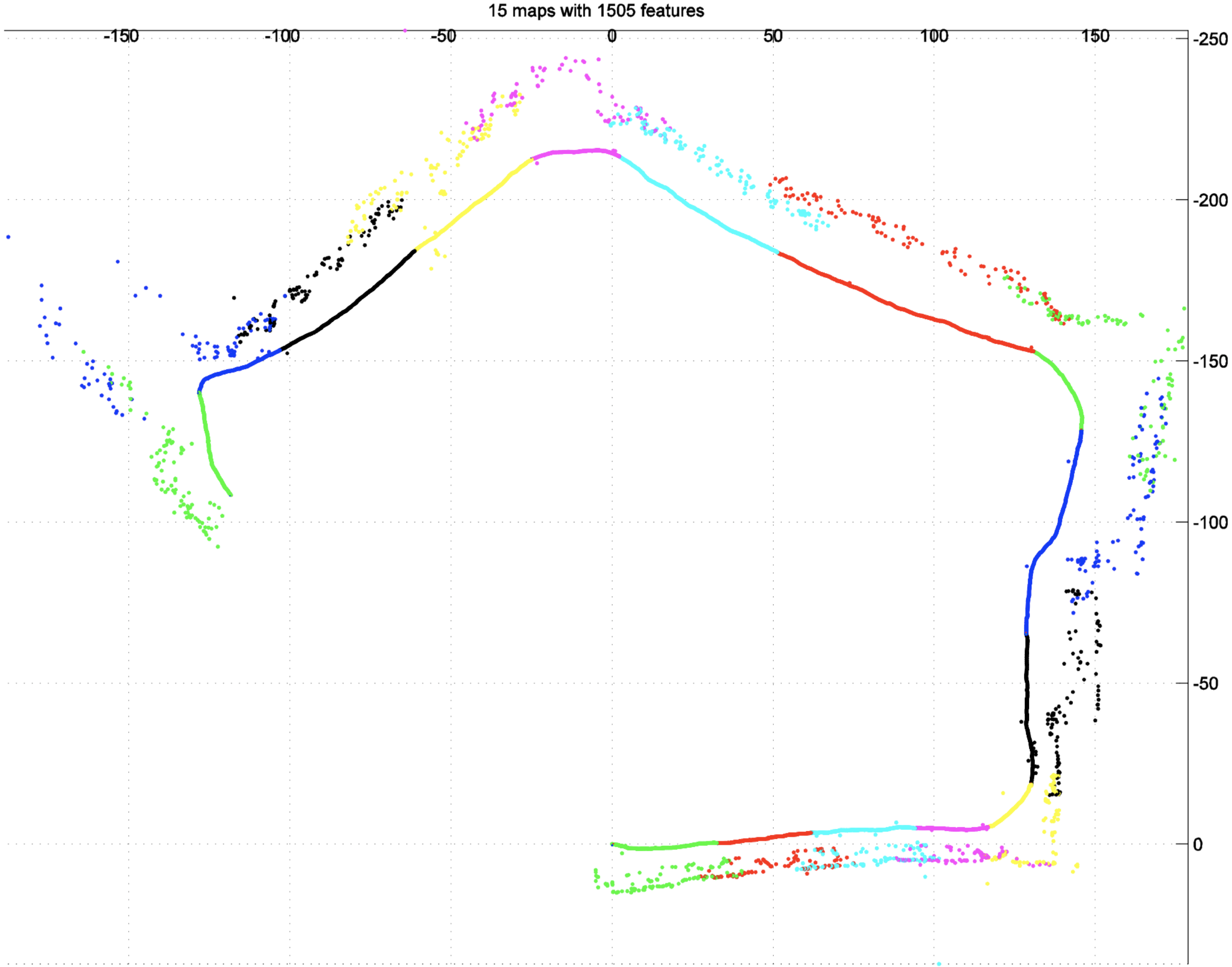


Sequence of local maps

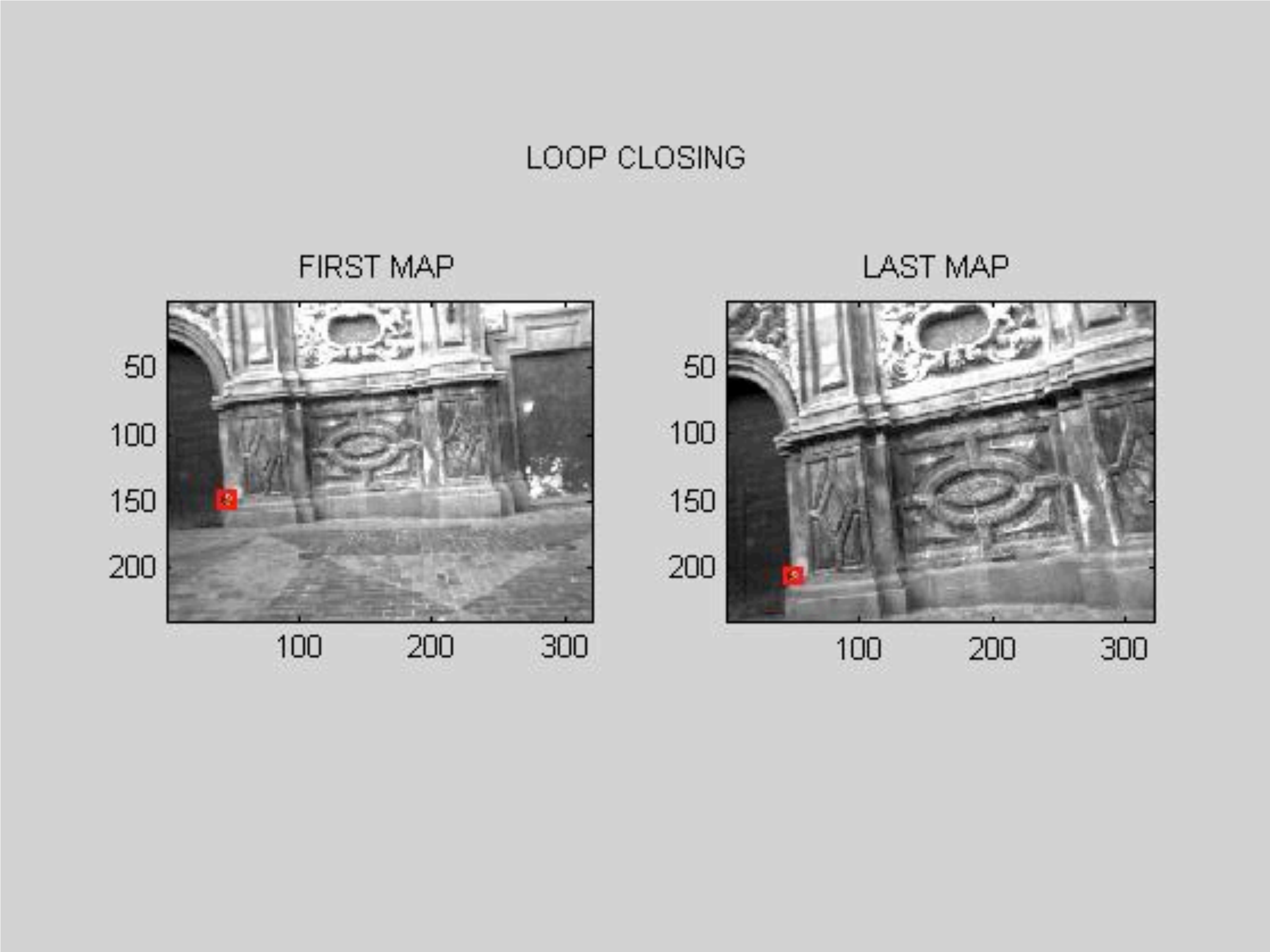


The scale is arbitrary (not observable)

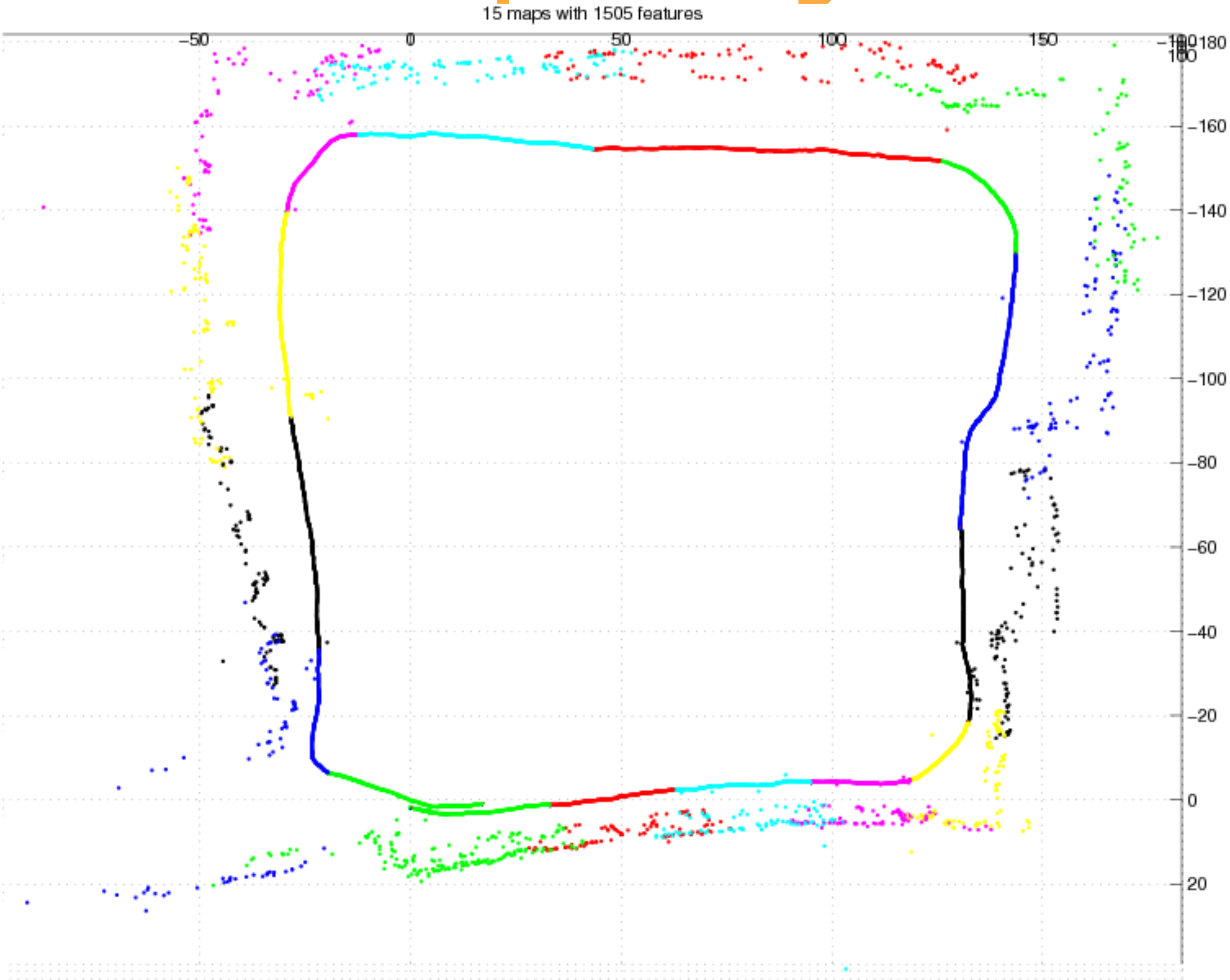
With scale compensation



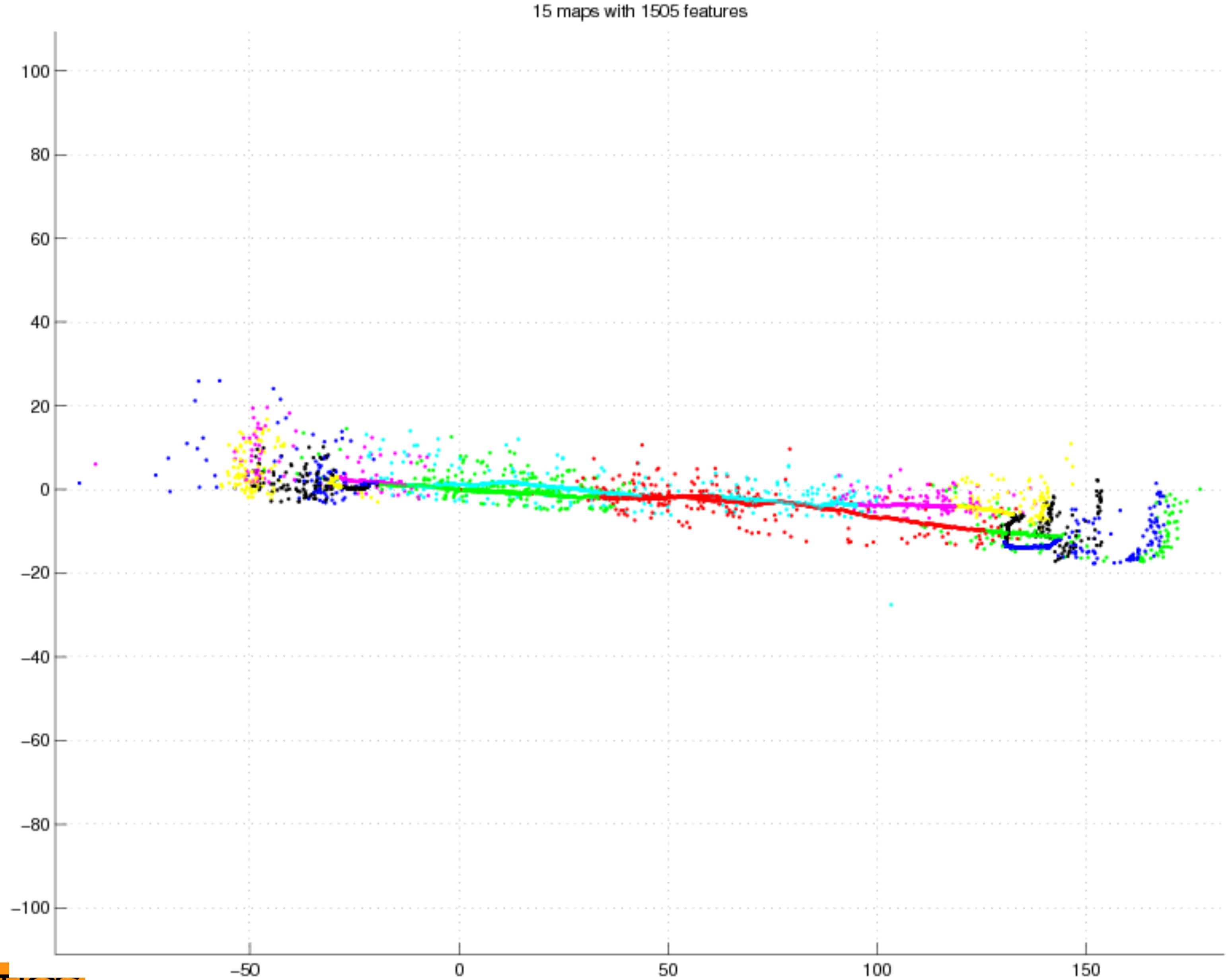
Loop closing: map-to-map matching



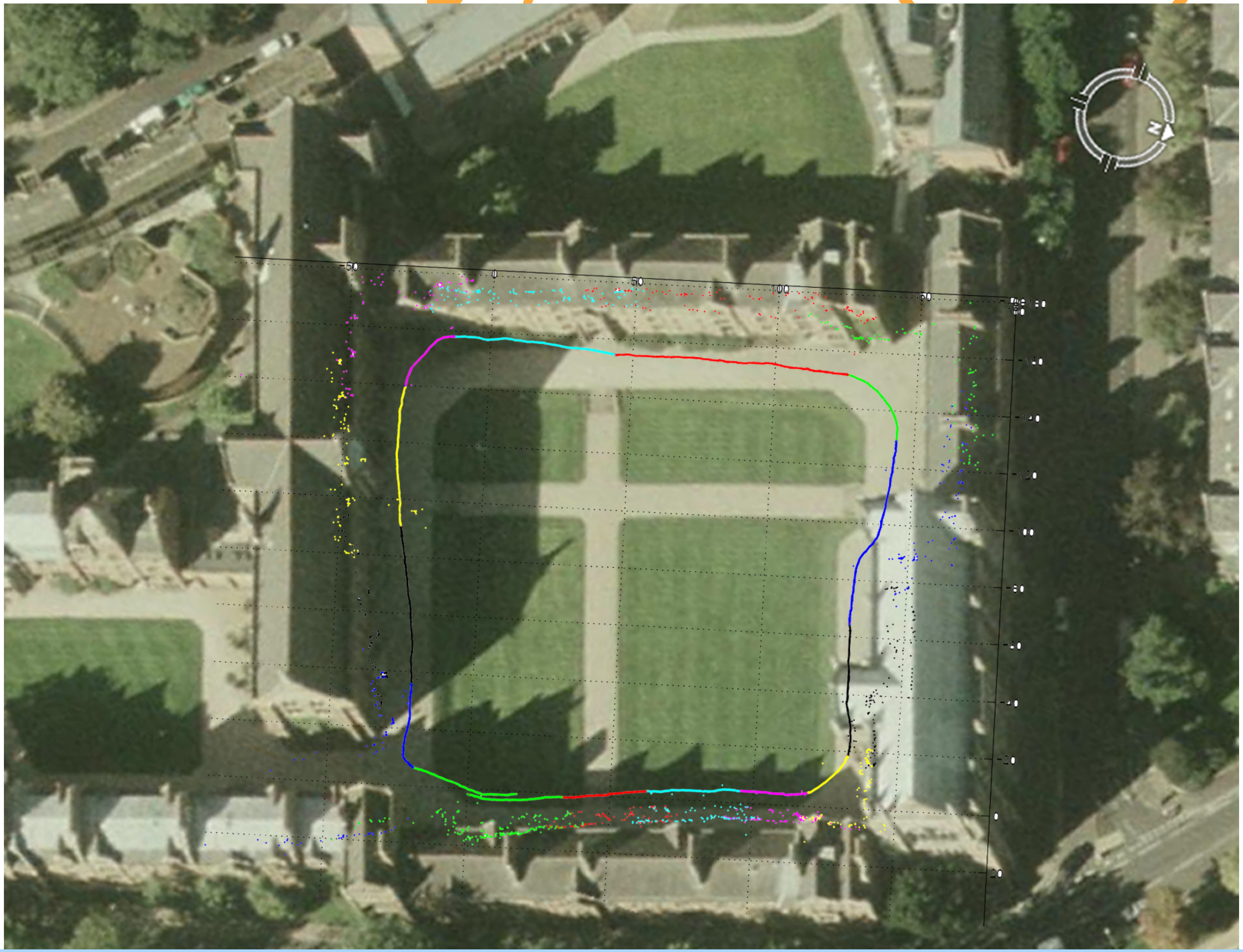
Loop closing



Loop closing (lateral view)



Keble College, Oxford (290m)



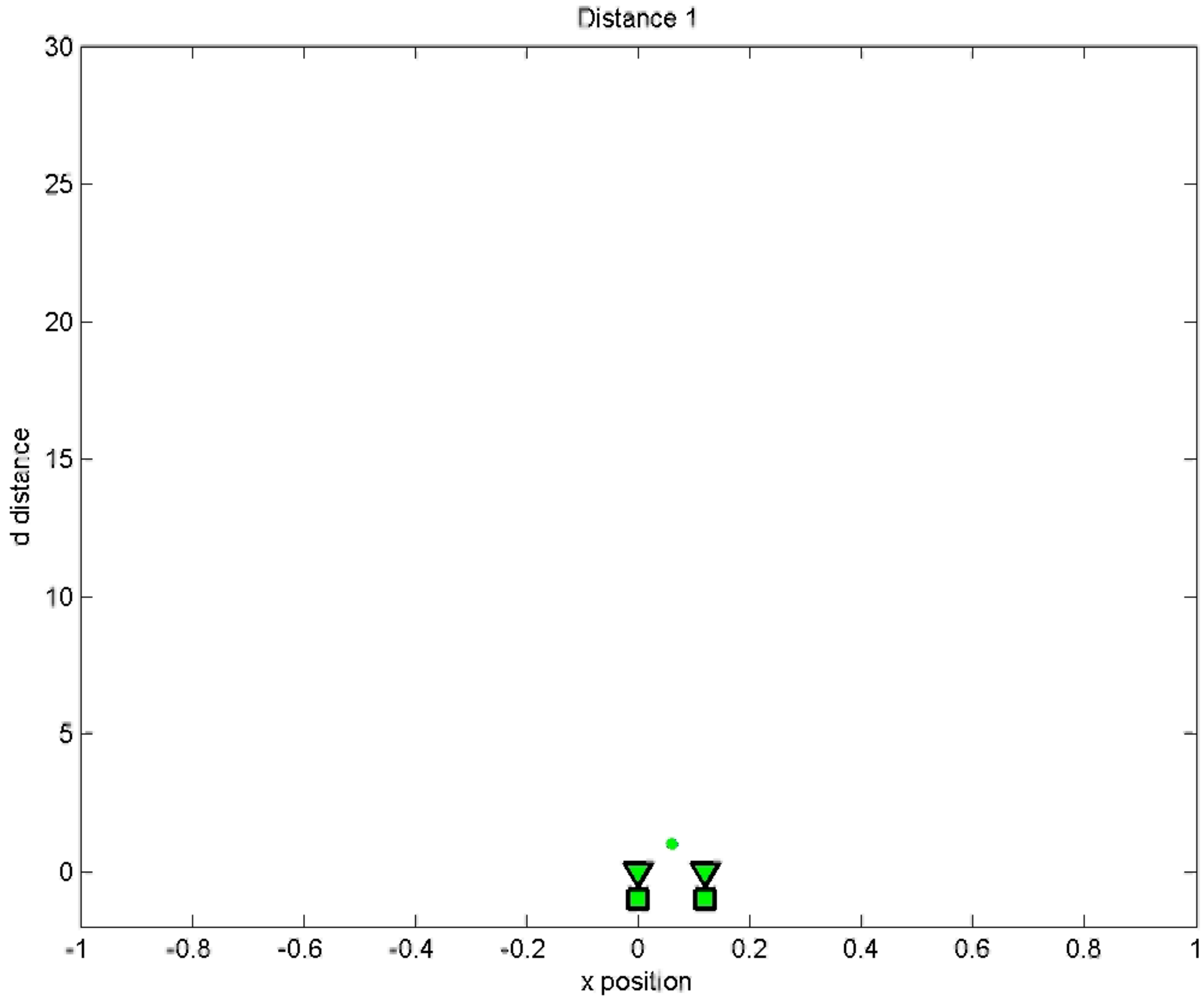
L. Clemente, A. Davison, I. Reid, J. Neira and J.D. Tardós **Mapping Large Loops with a Single Hand-Held Camera**. Robotics: Science and Systems, 2007.

SLAM using only stereo

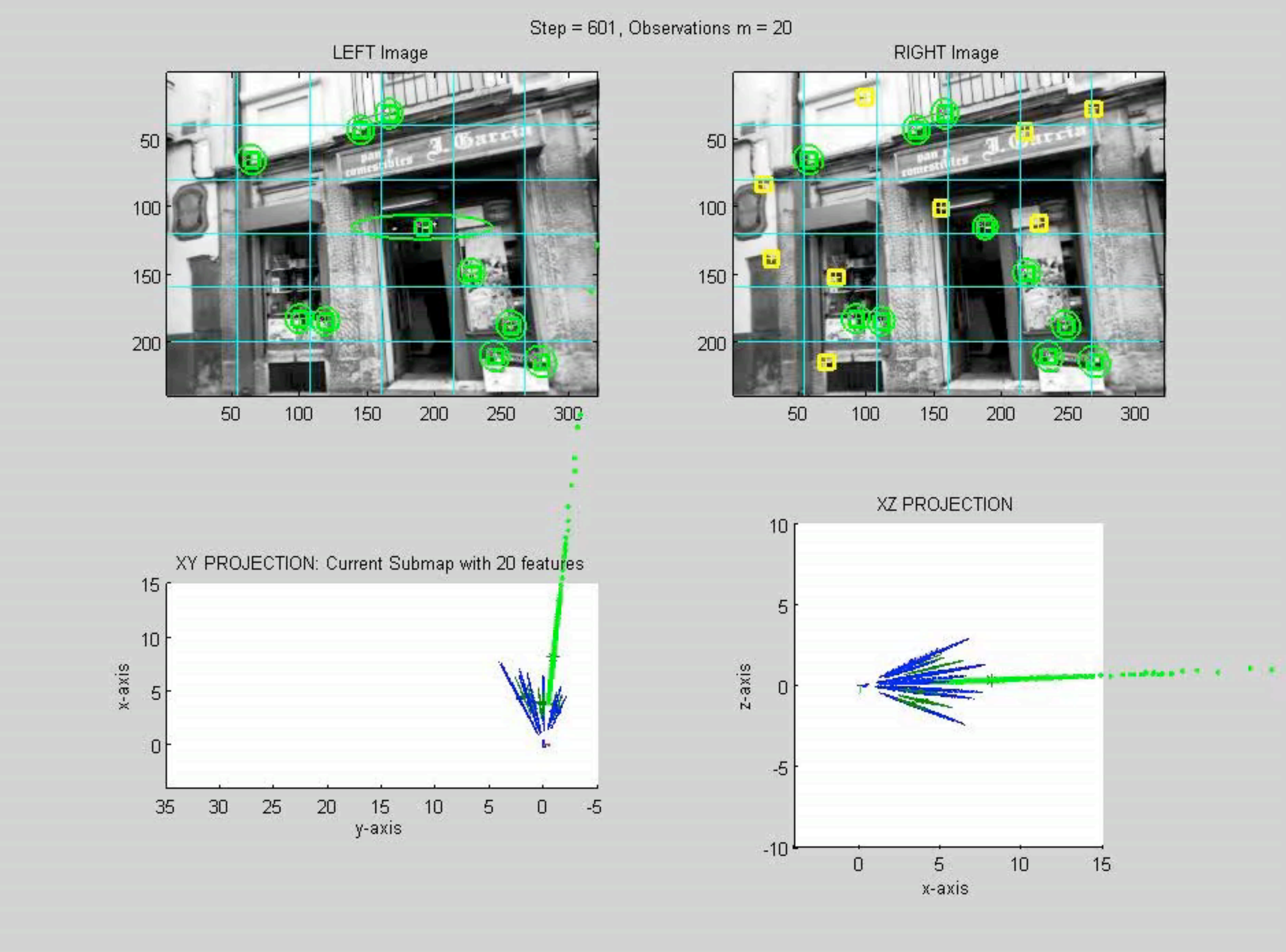
- Experimental setup

A bumblebee, a laptop and a firewire cable

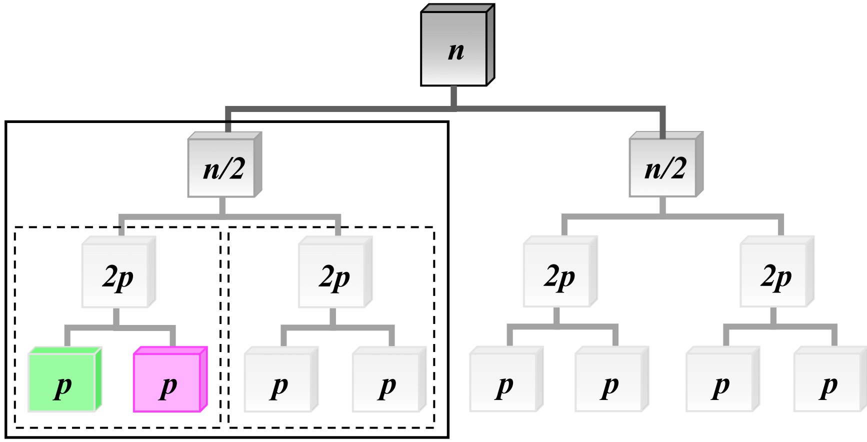
Depth .vs. Inverse Depth

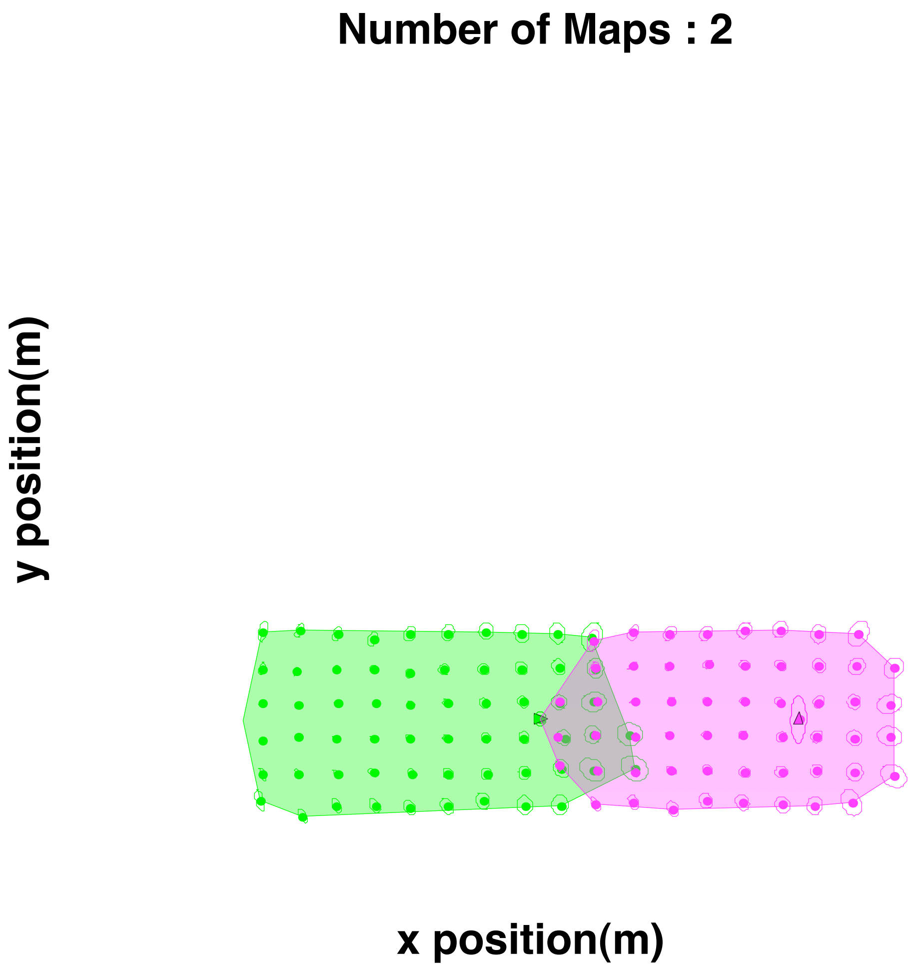


Basic EKF SLAM

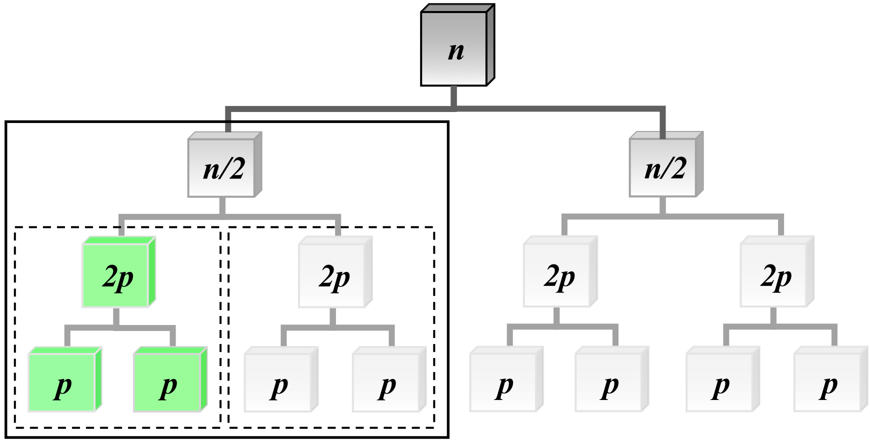


Divide & Conquer SLAM

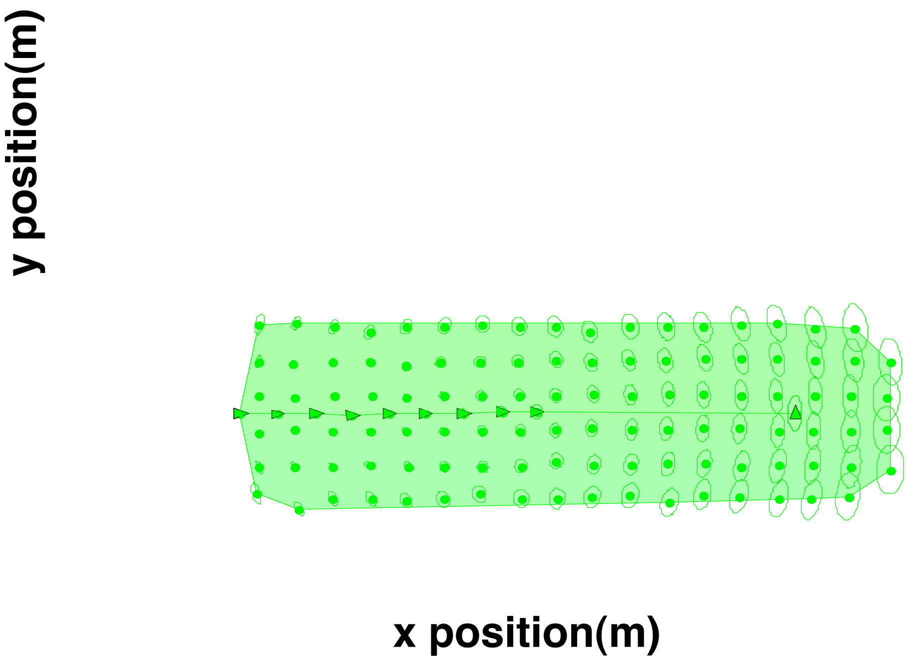




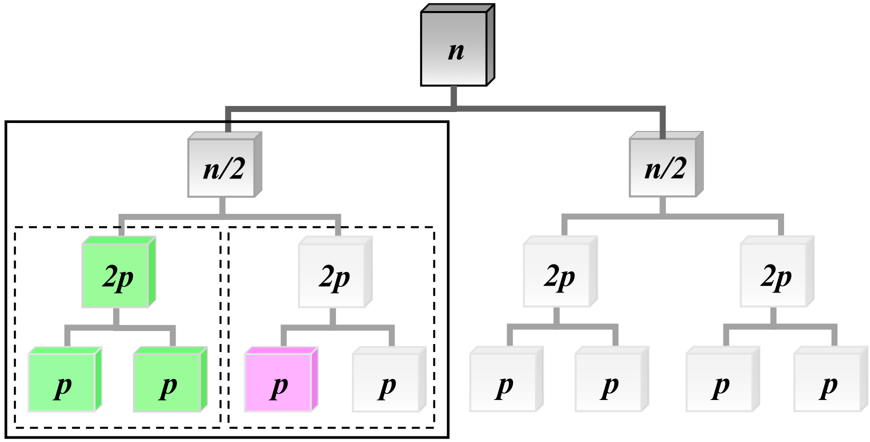
Divide & Conquer SLAM

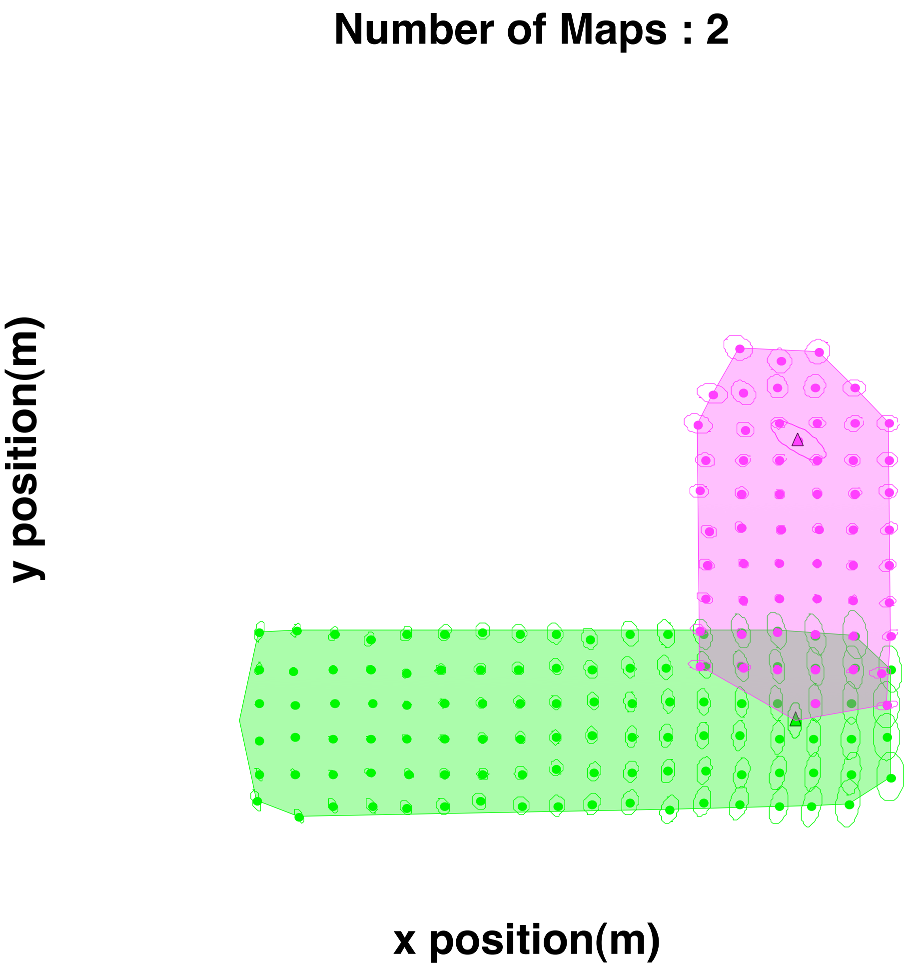


Number of Maps : 1

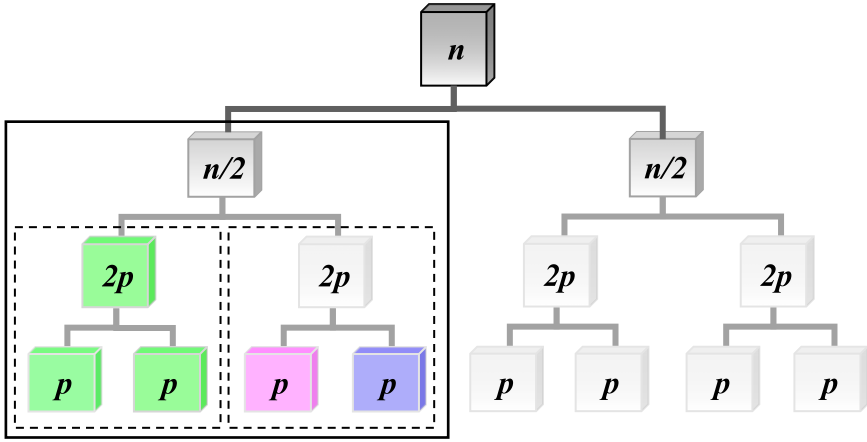


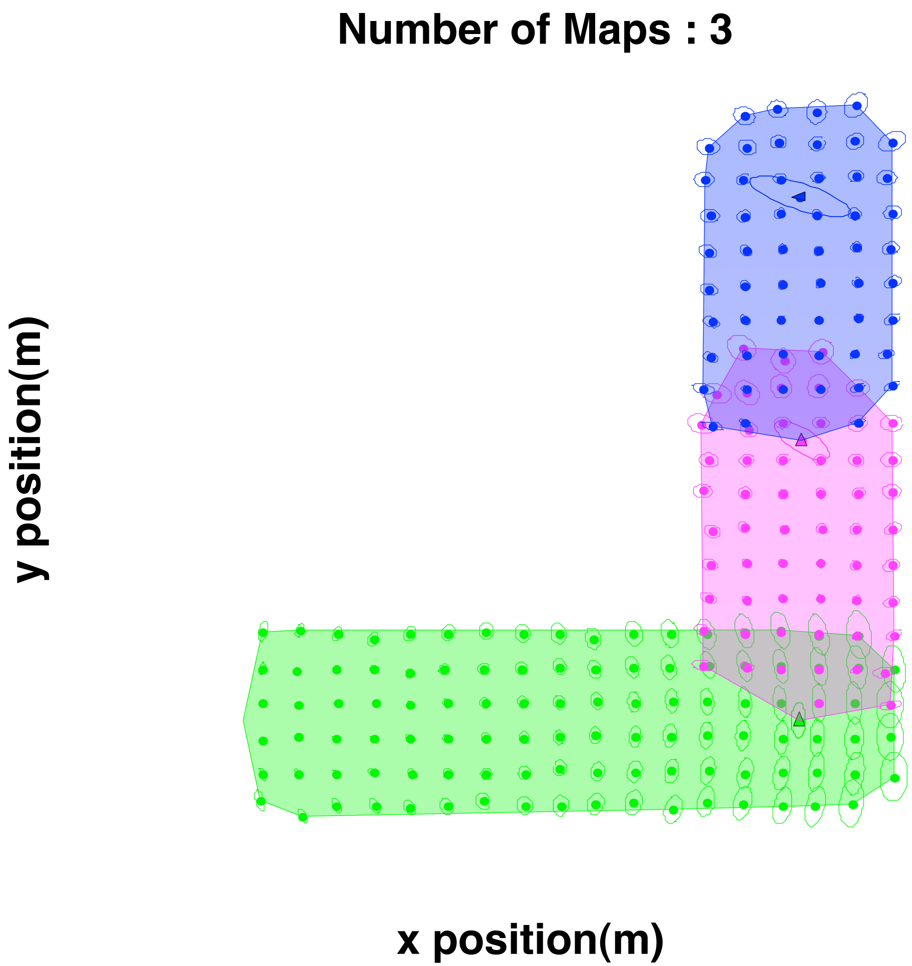
Divide & Conquer SLAM



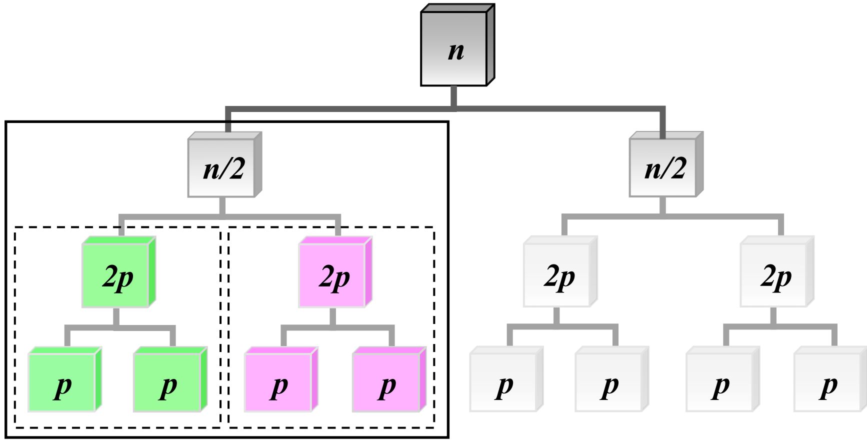


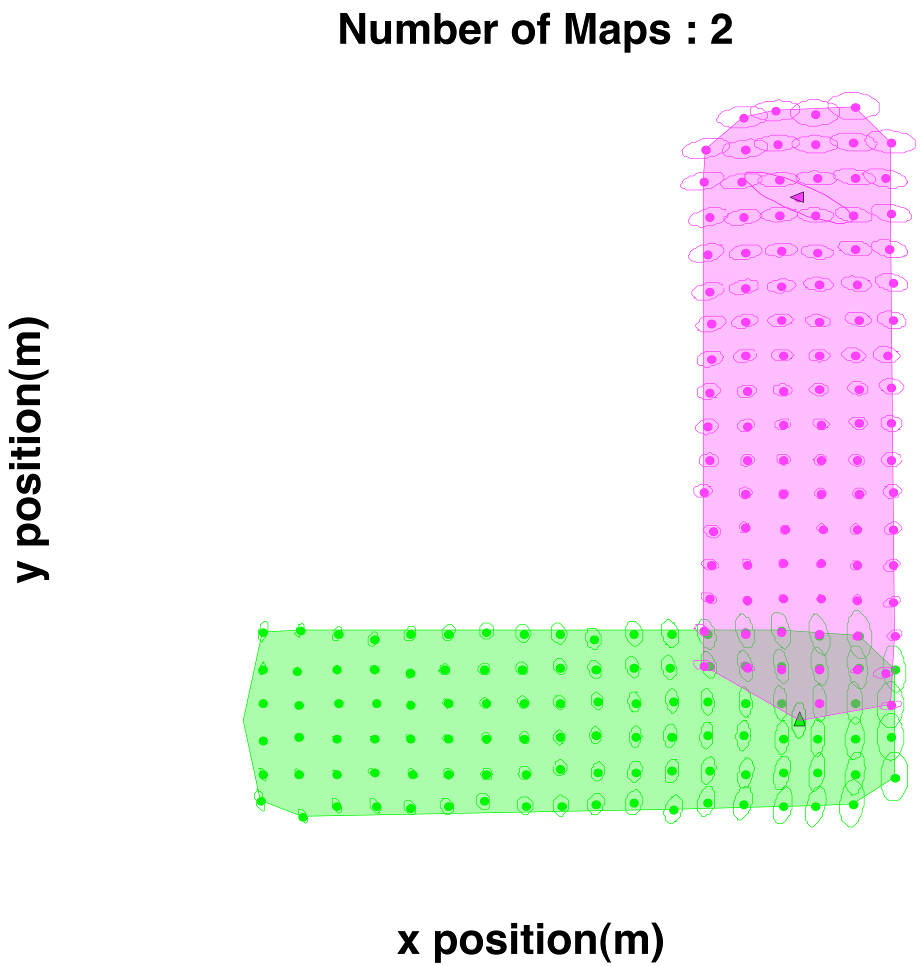
Divide & Conquer SLAM



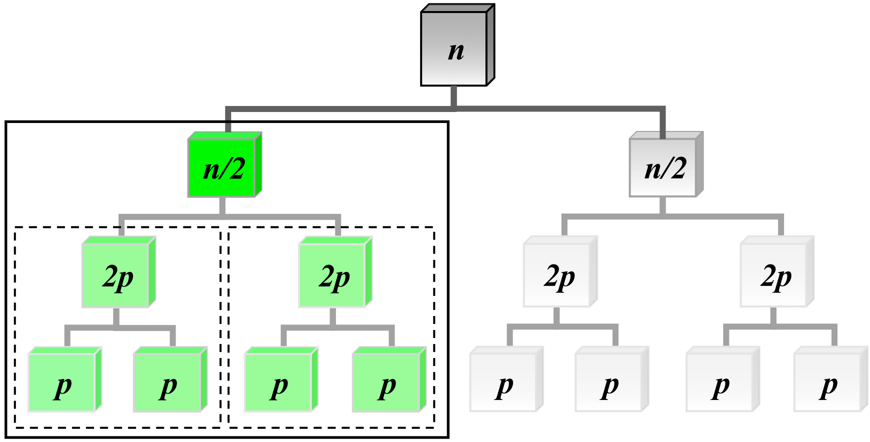


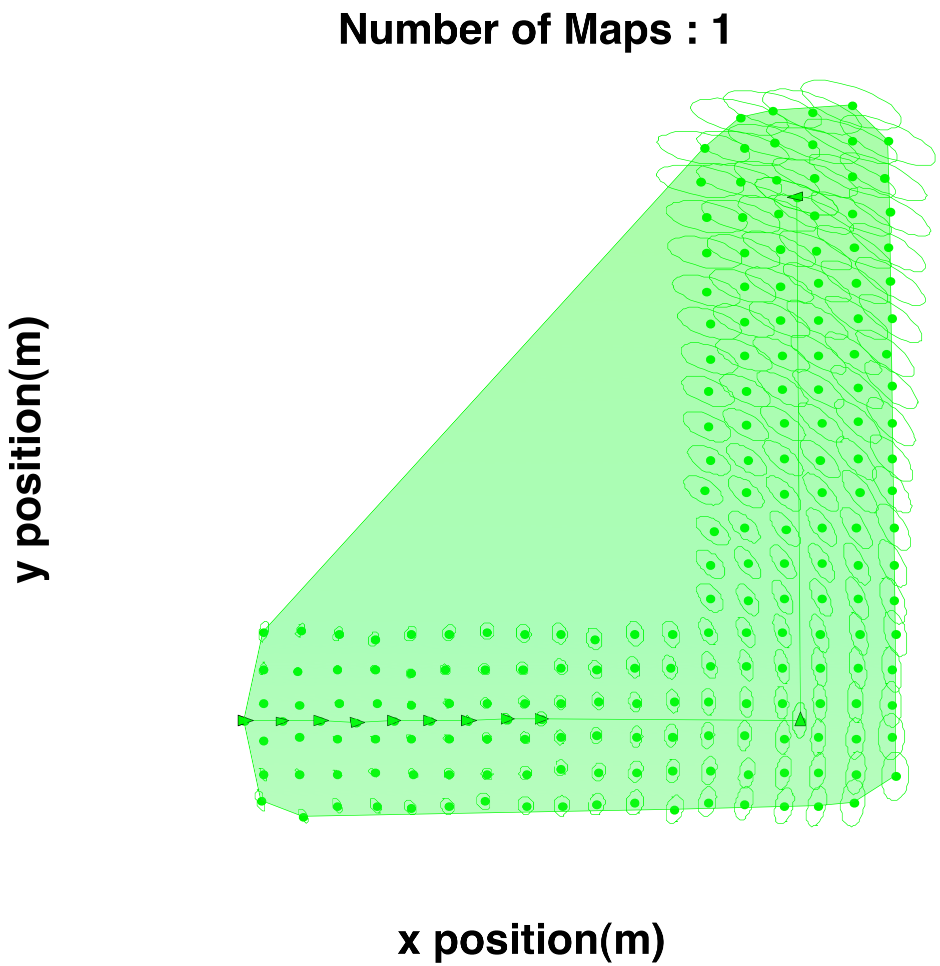
Divide & Conquer SLAM



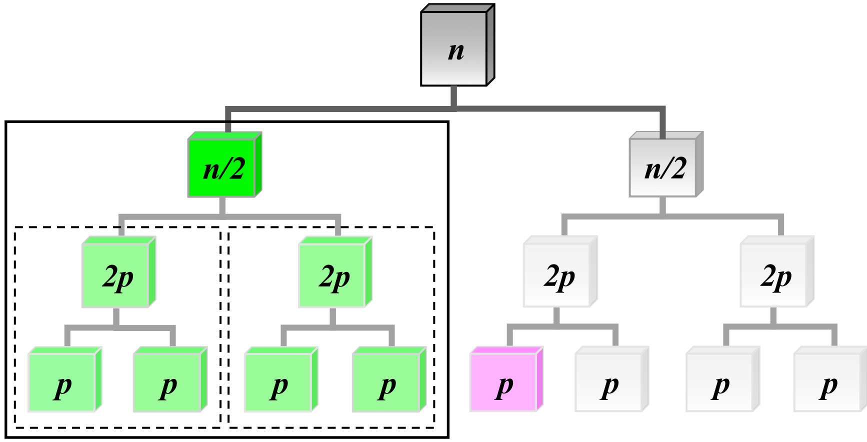


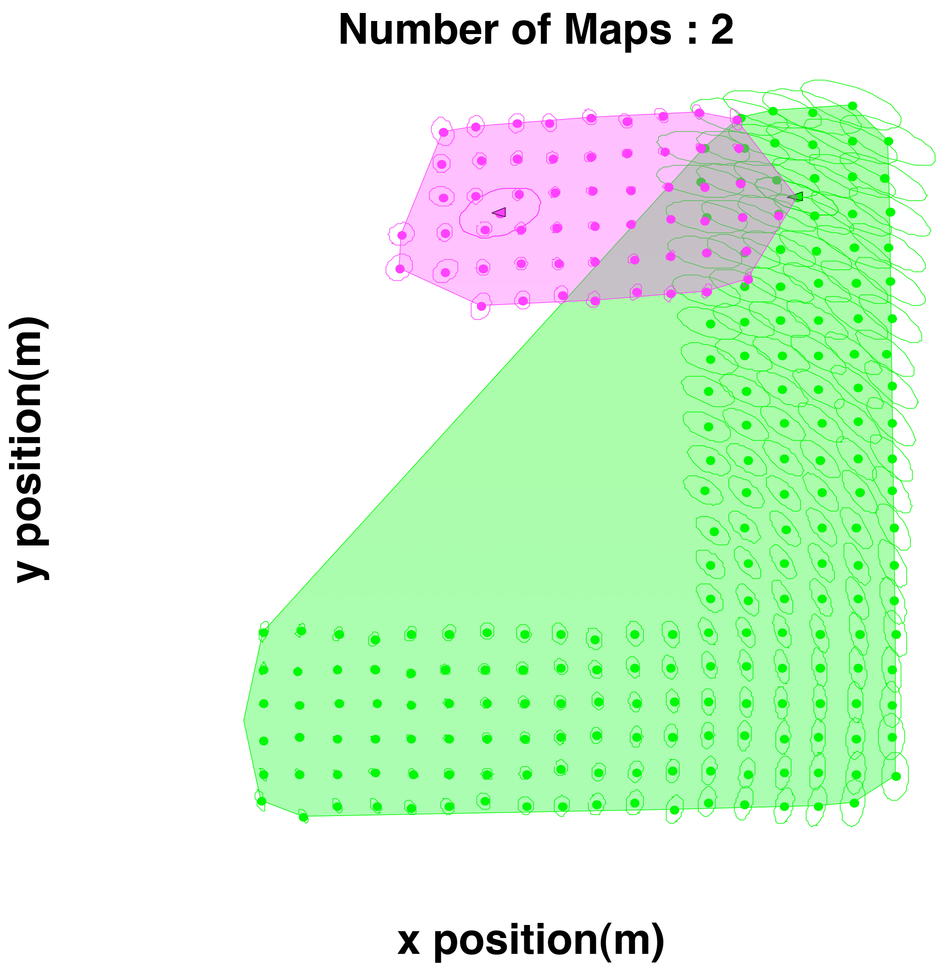
Divide & Conquer SLAM



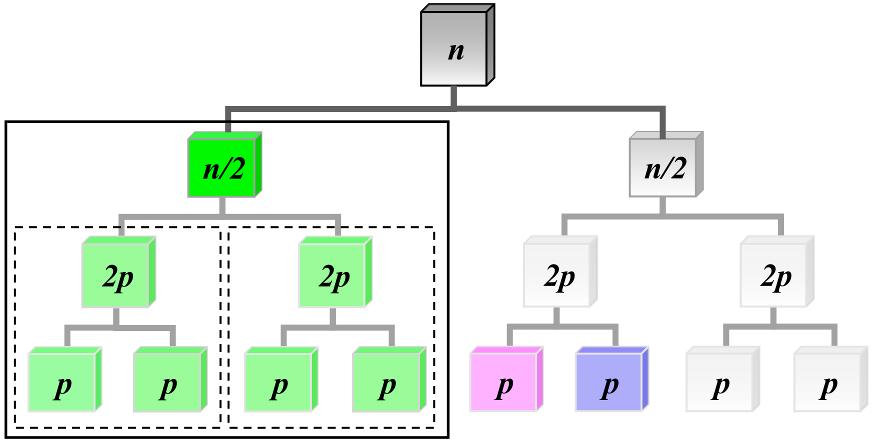


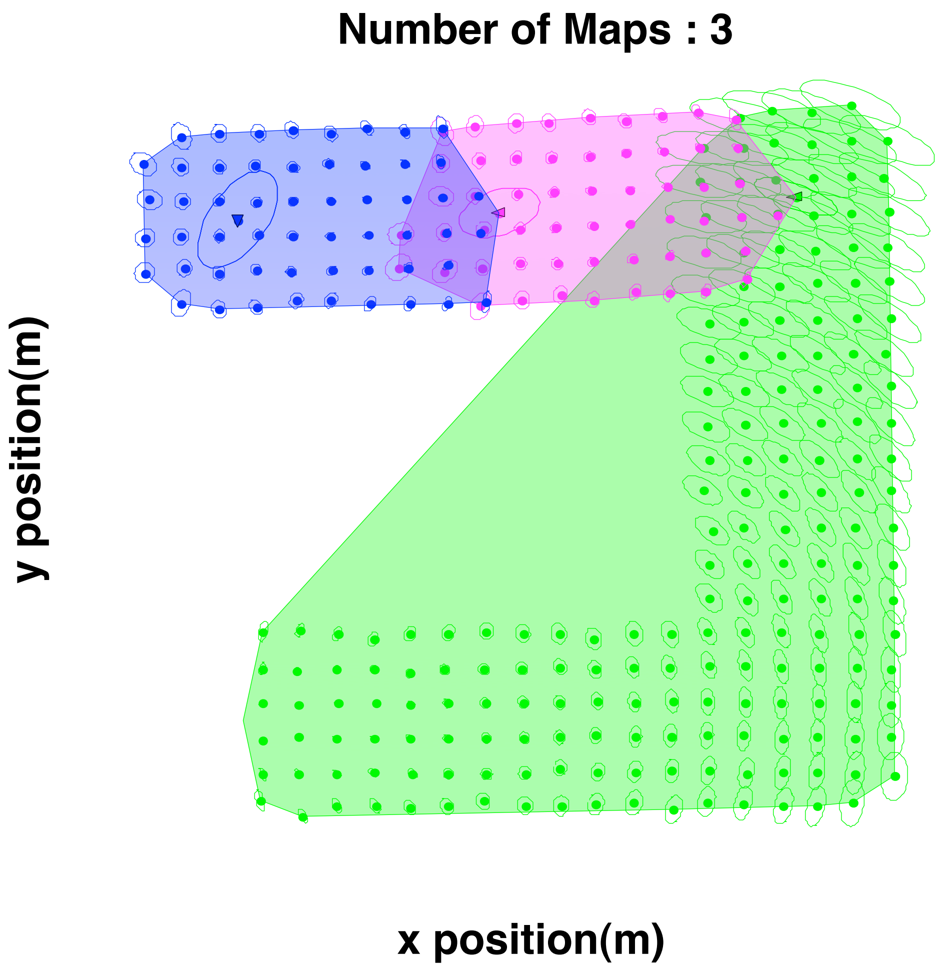
Divide & Conquer SLAM



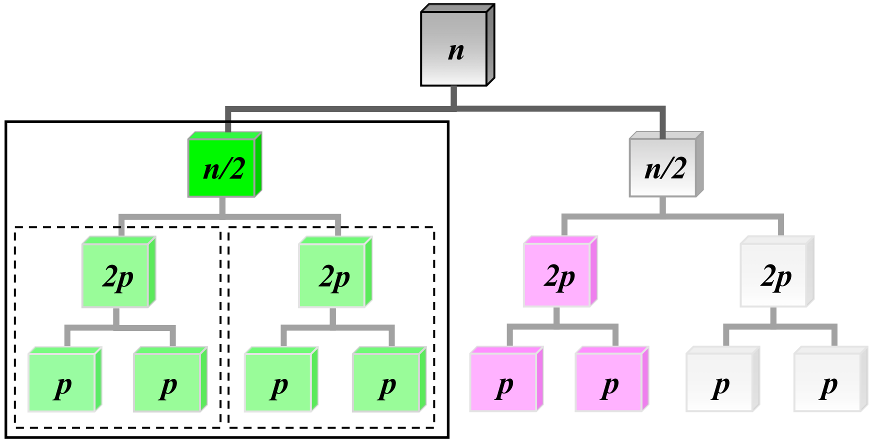


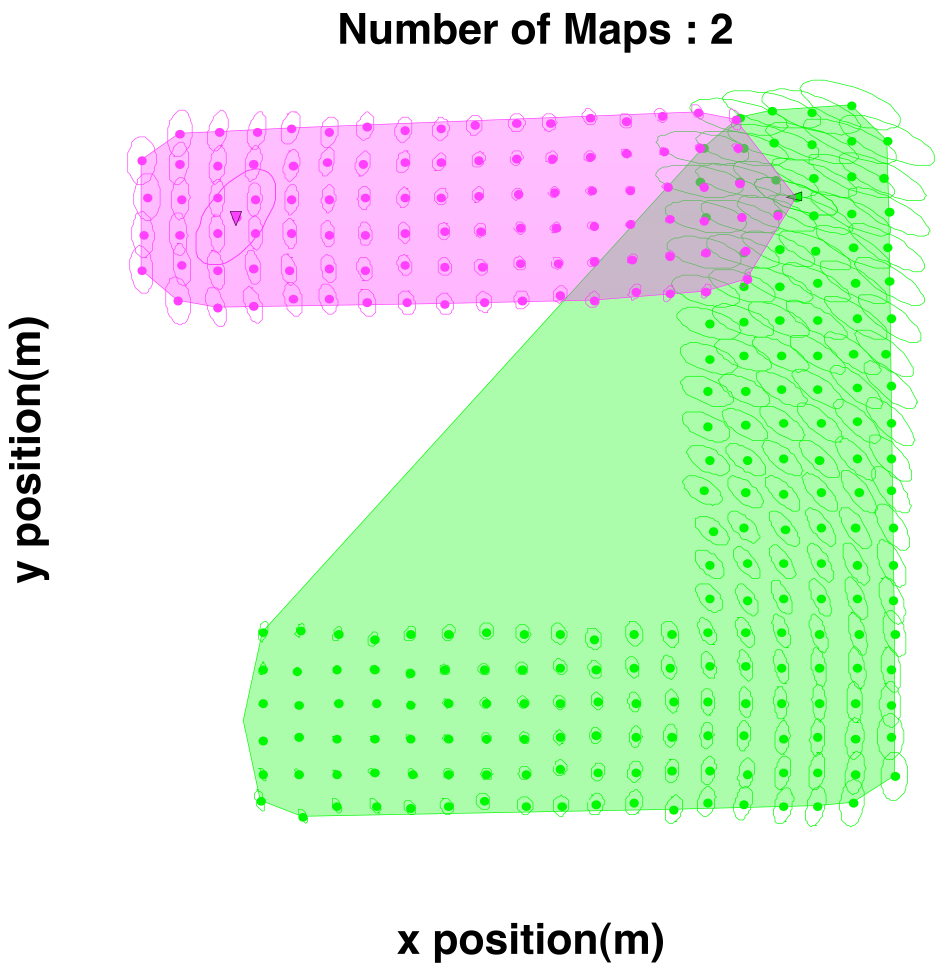
Divide & Conquer SLAM



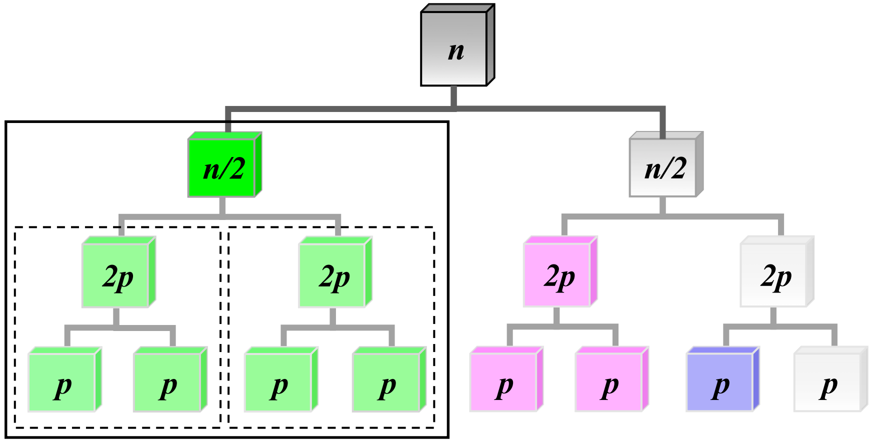


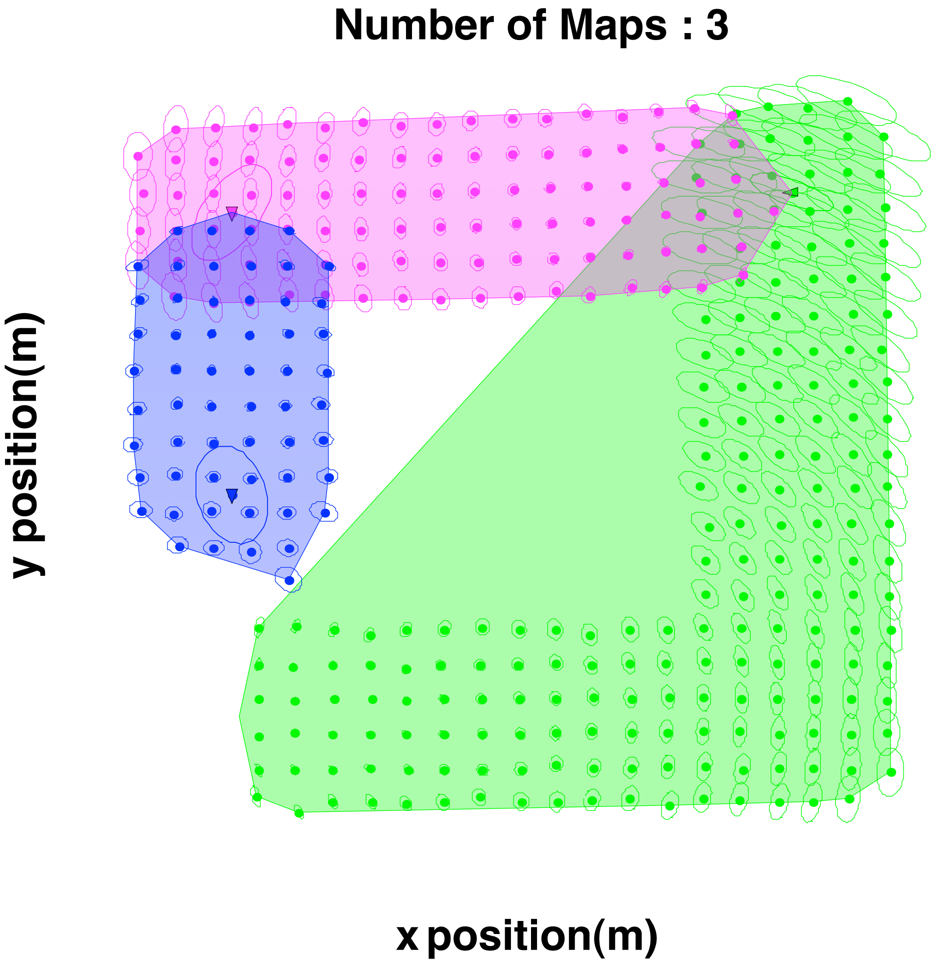
Divide & Conquer SLAM



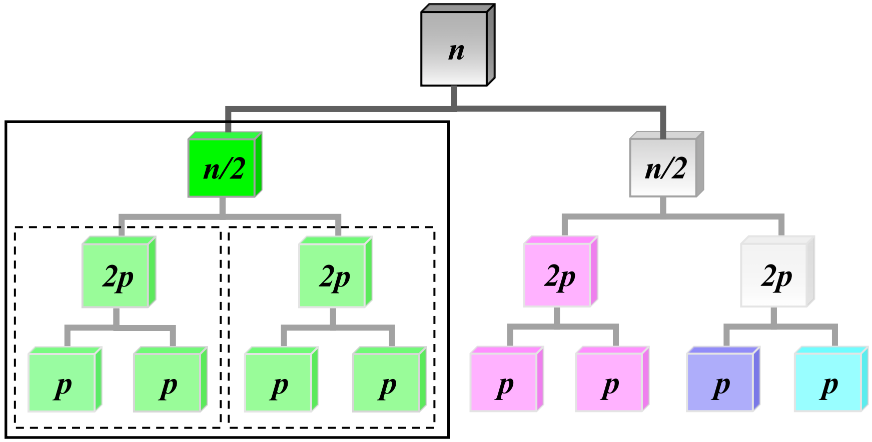


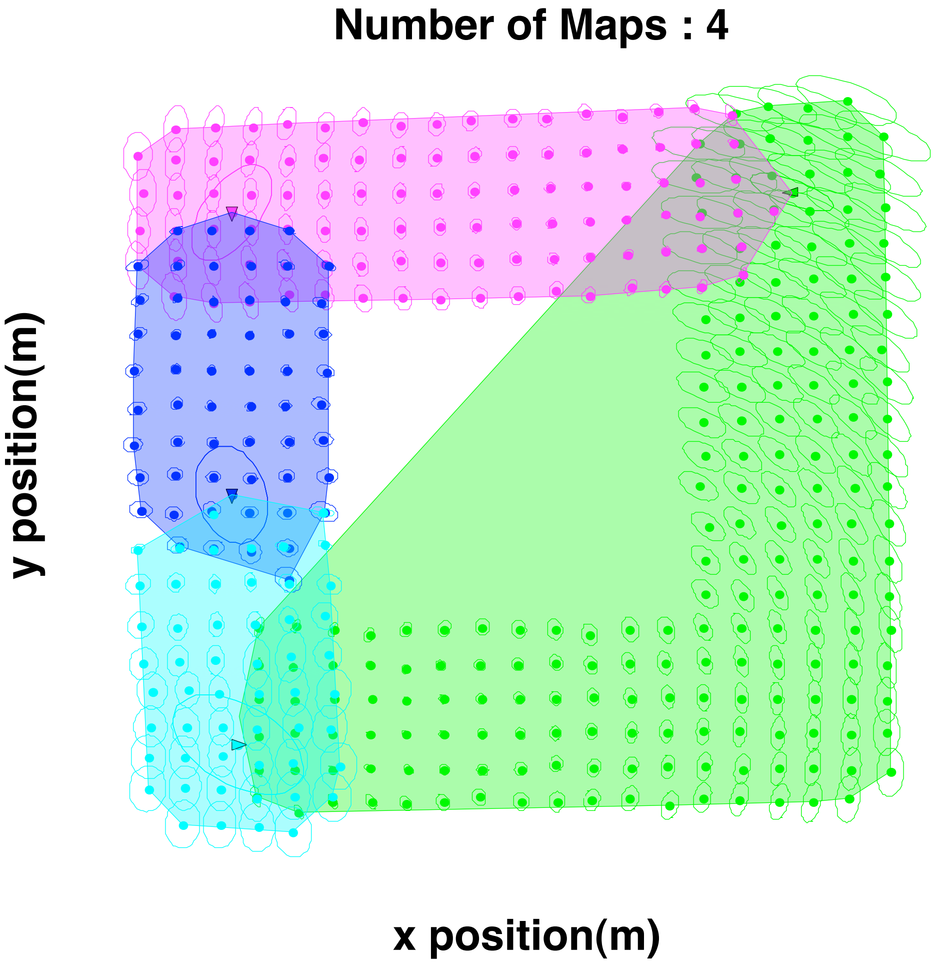
Divide & Conquer SLAM



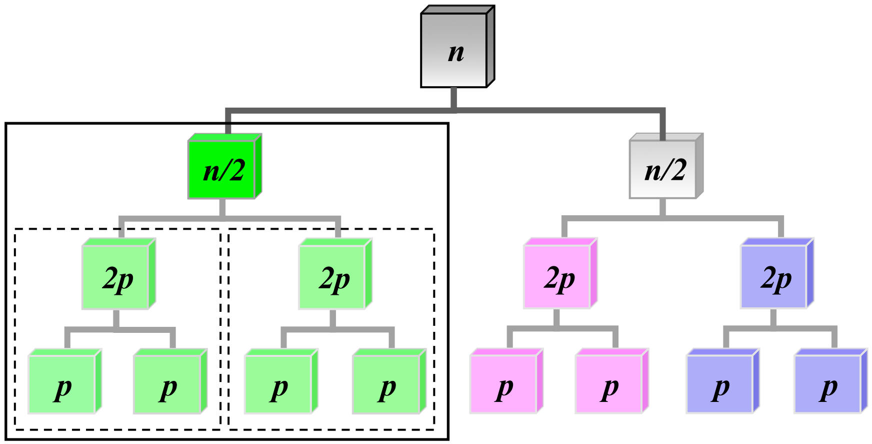


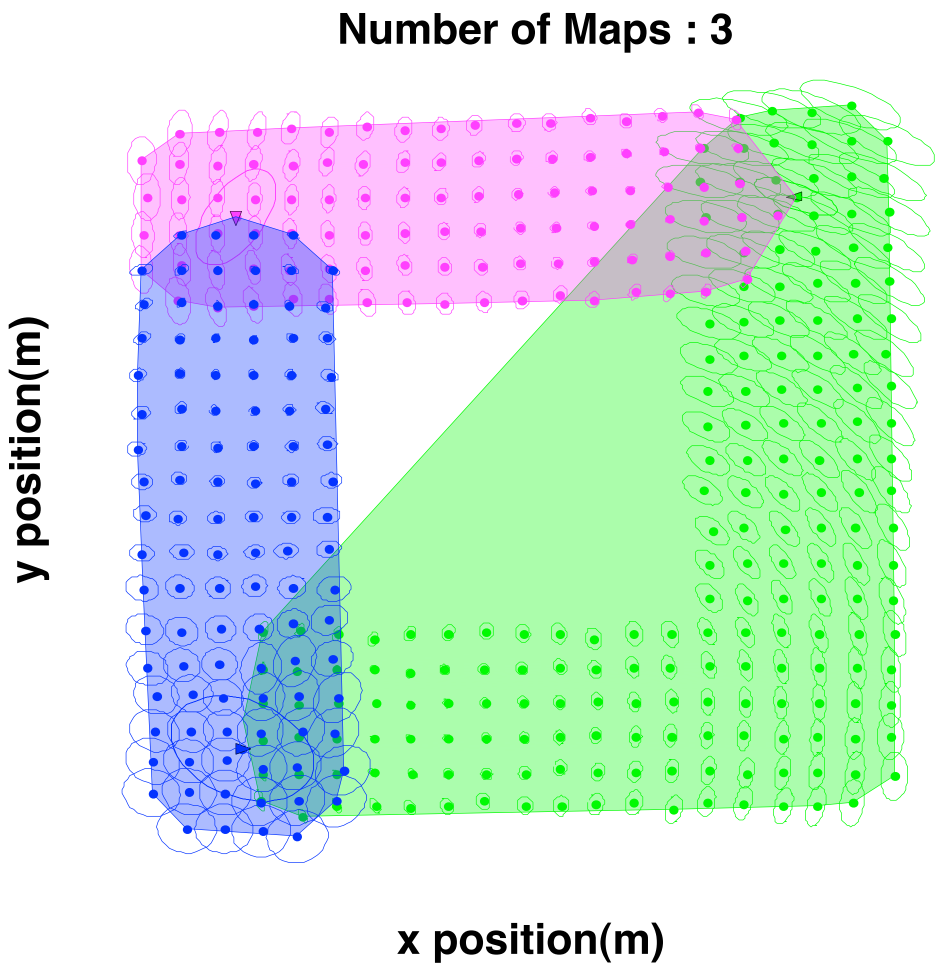
Divide & Conquer SLAM



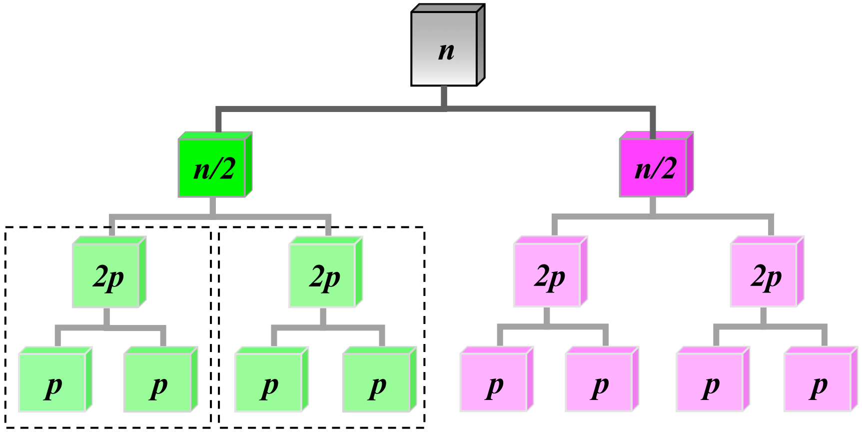


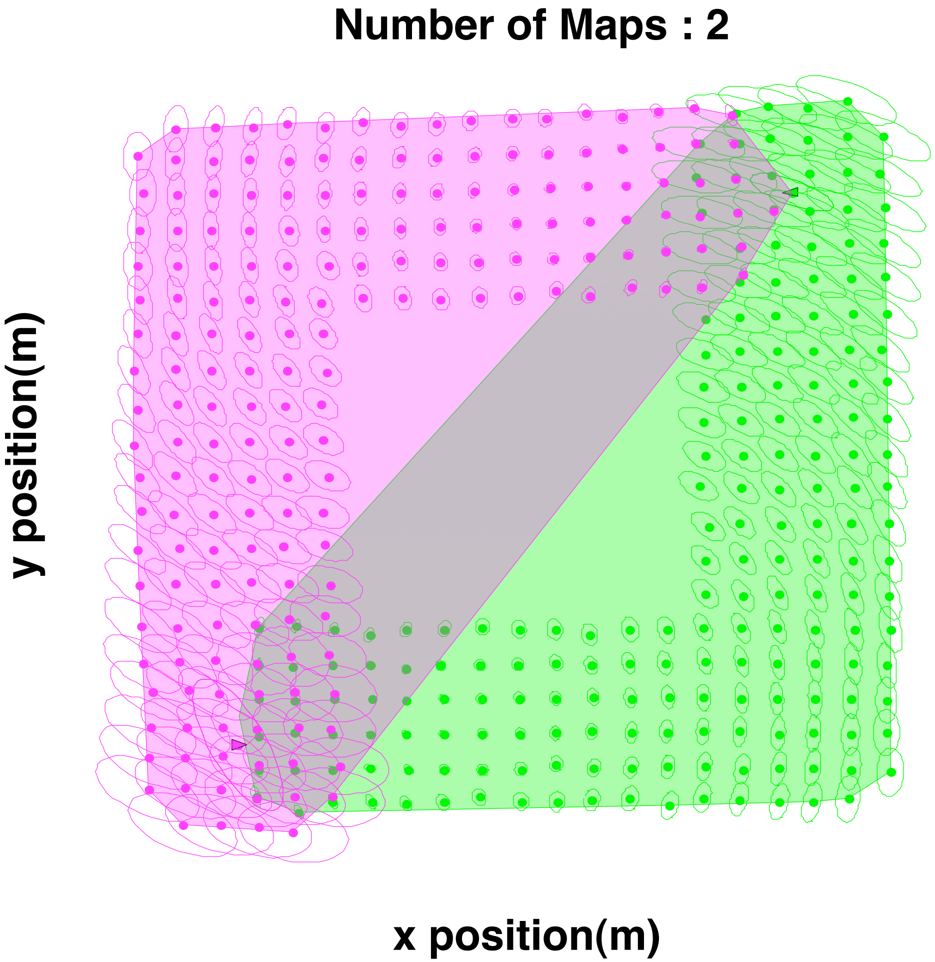
Divide & Conquer SLAM



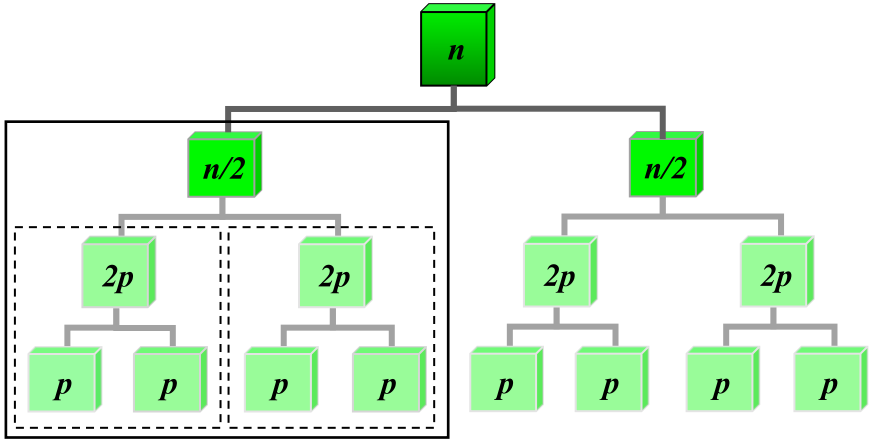


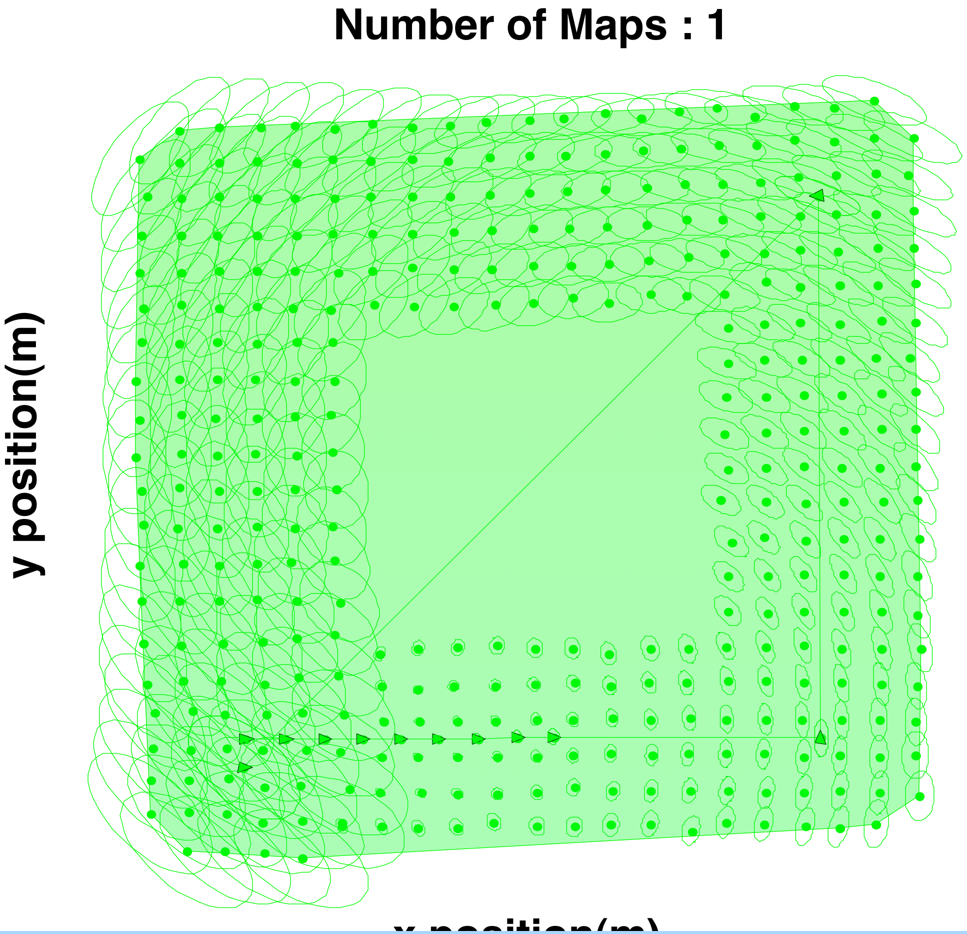
Divide & Conquer SLAM





Divide & Conquer SLAM

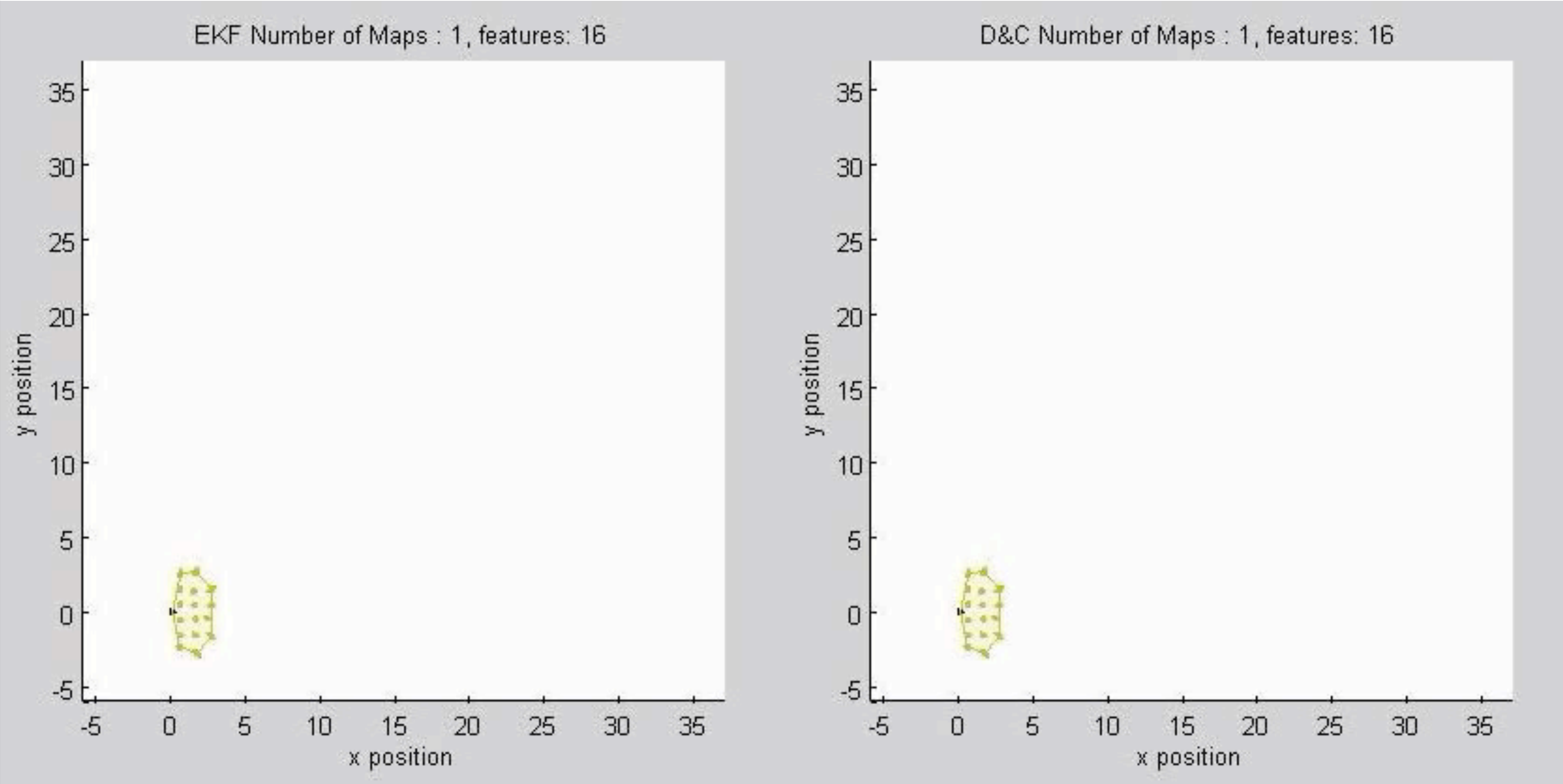




L.M. Paz, P. Jensfelt, J.D. Tardós and J. Neira. **EKF SLAM updates in O(n) with Divide and Conquer SLAM** 2007 IEEE Int. Conf. Robotics and Automation, April 10-14, Rome, Italy

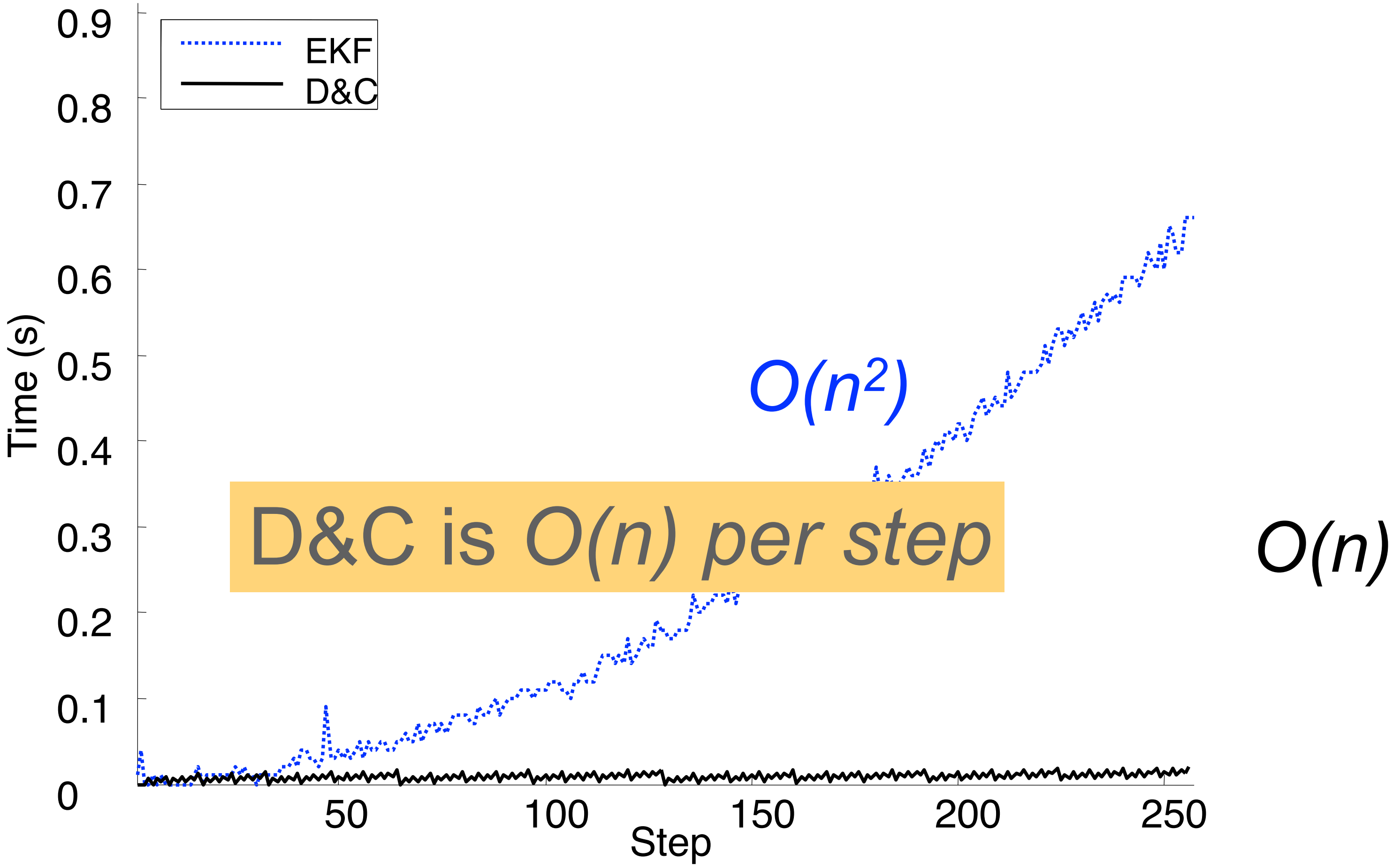
ROBOTICS UNIVERSIDAD DE ZARAGOZA

Loop Trajectory



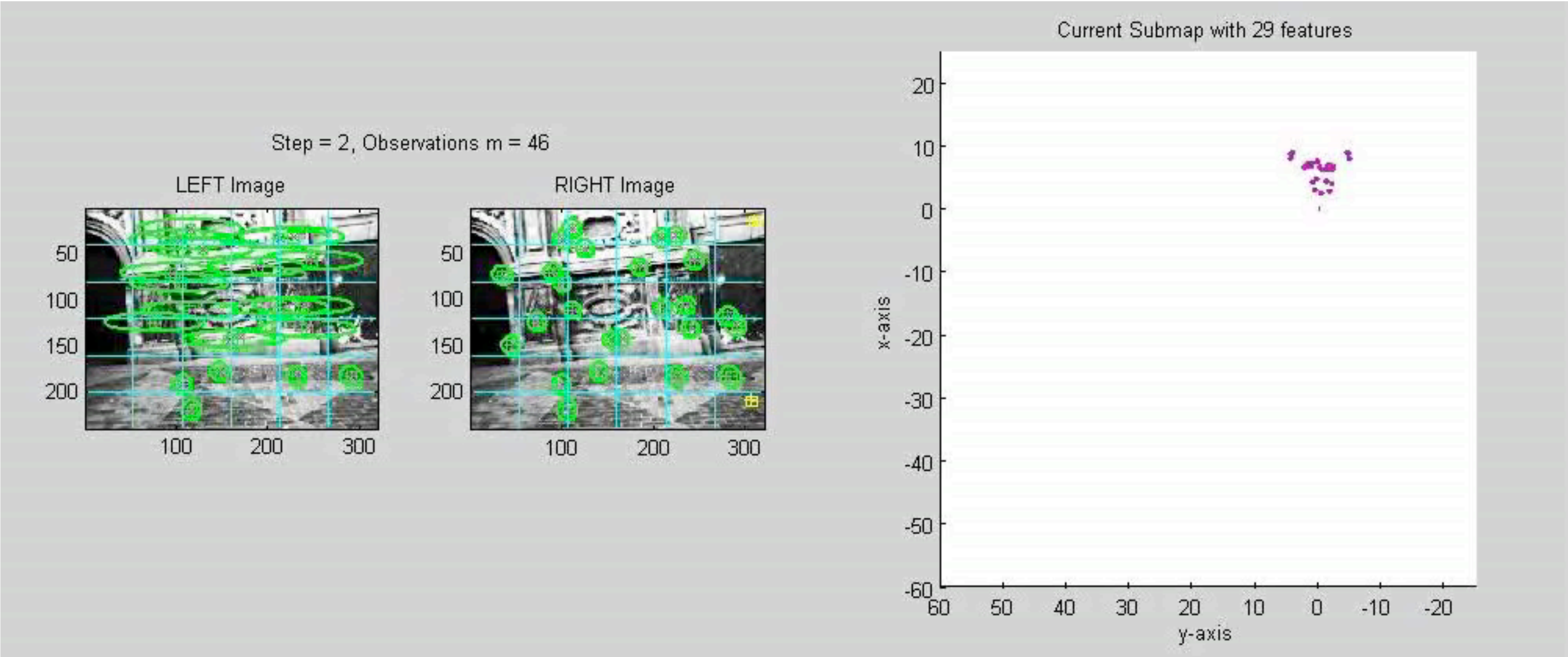
L. Paz, J. Neira and J.D. Tardós **Divide and Conquer: EKF SLAM in O(n)**. IEEE Transactions on Robotics, October 2008.

Amortized cost per step



D&C is *O(n) per step*

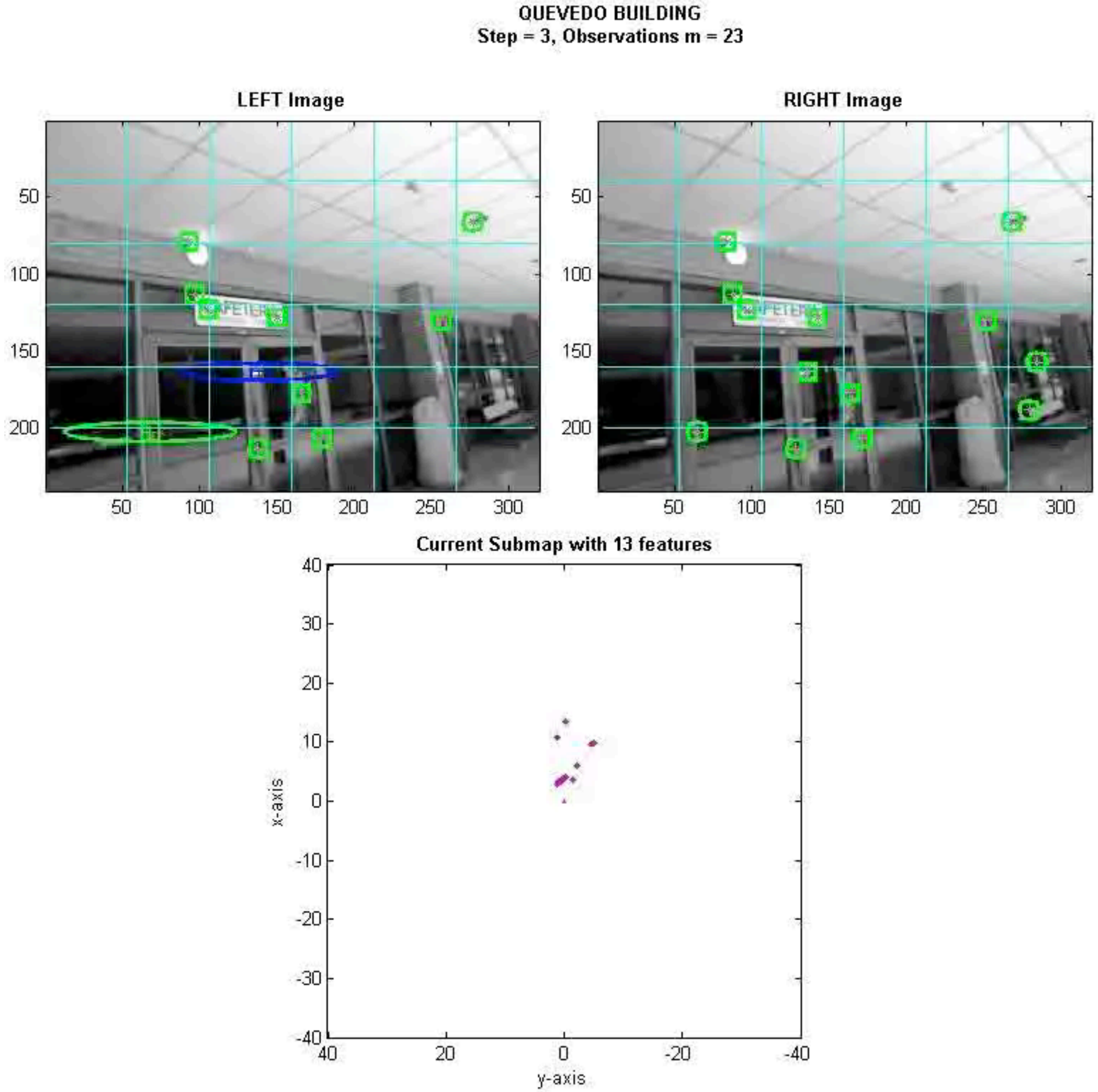
6DOF SLAM with stereo



L. Paz, P. Piníes, J. Neira and J.D. Tardós **Large Scale 6DOF SLAM with Stereo-in-Hand.** IEEE Transactions on Robotics, 2008.

6Dof Stereo SLAM, outdoors

6Dof Stereo SLAM, indoors



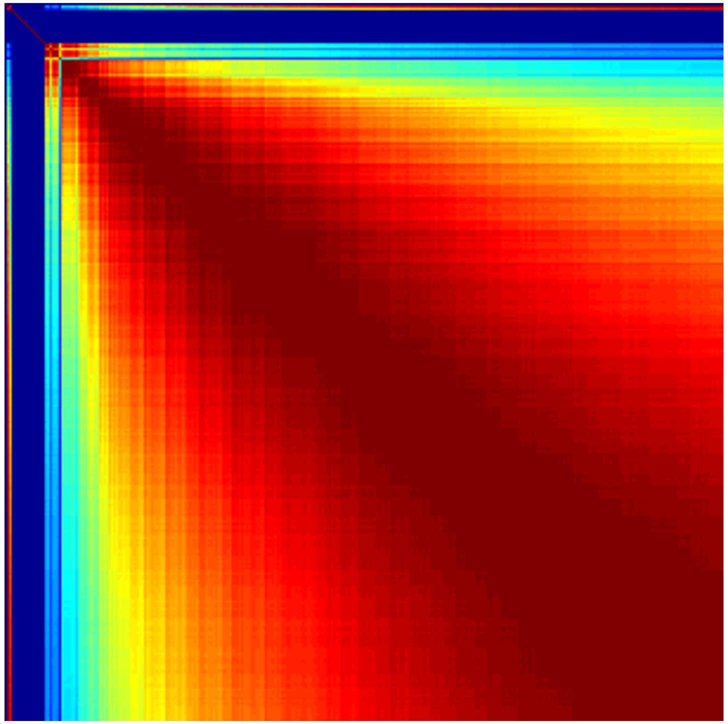
6Dof Stereo SLAM, indoors

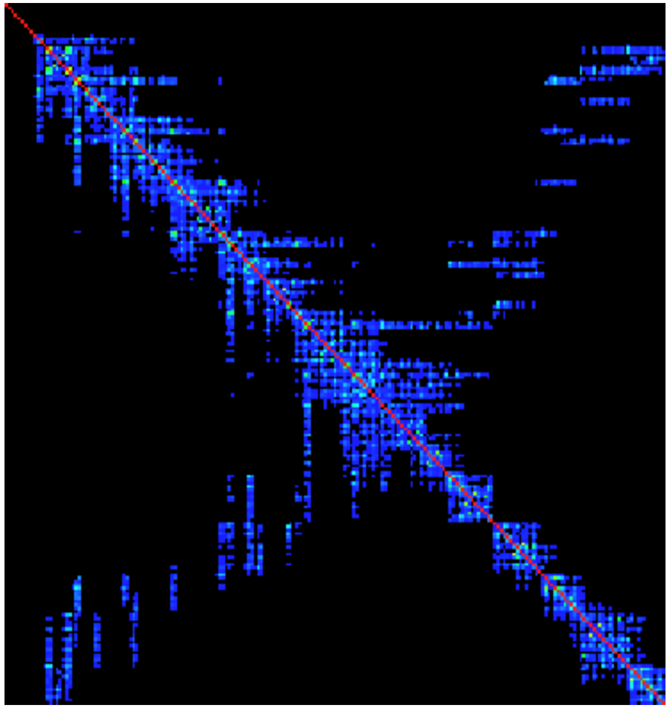
Technische Universität München, 2010

Scaling: information filters

- Covariance Matrices are full
- Joins are full

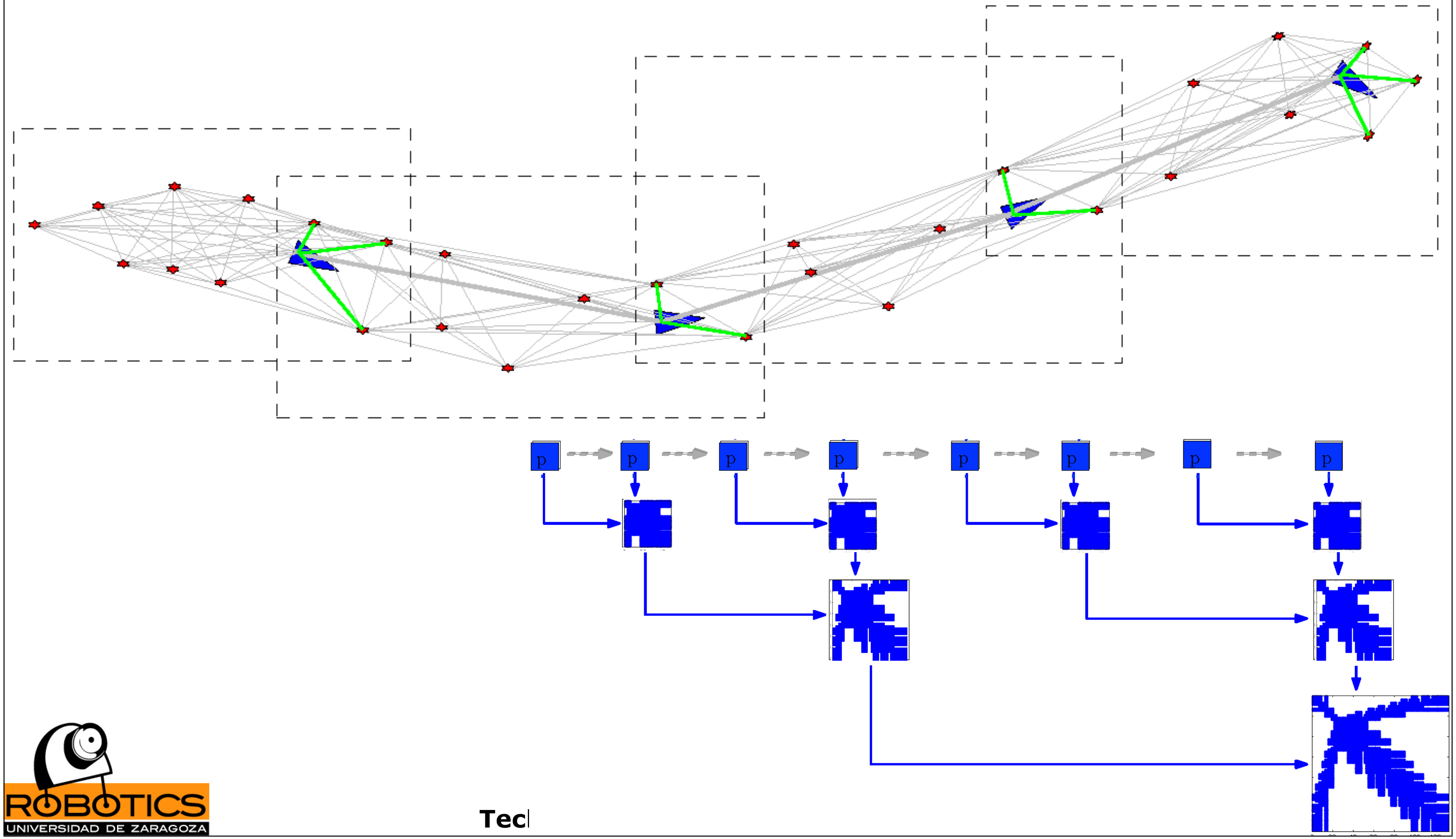
- Information Matrices are app. sparse (Thrun, 2006)
- Joins are **exactly** sparse (Huang, 2008)





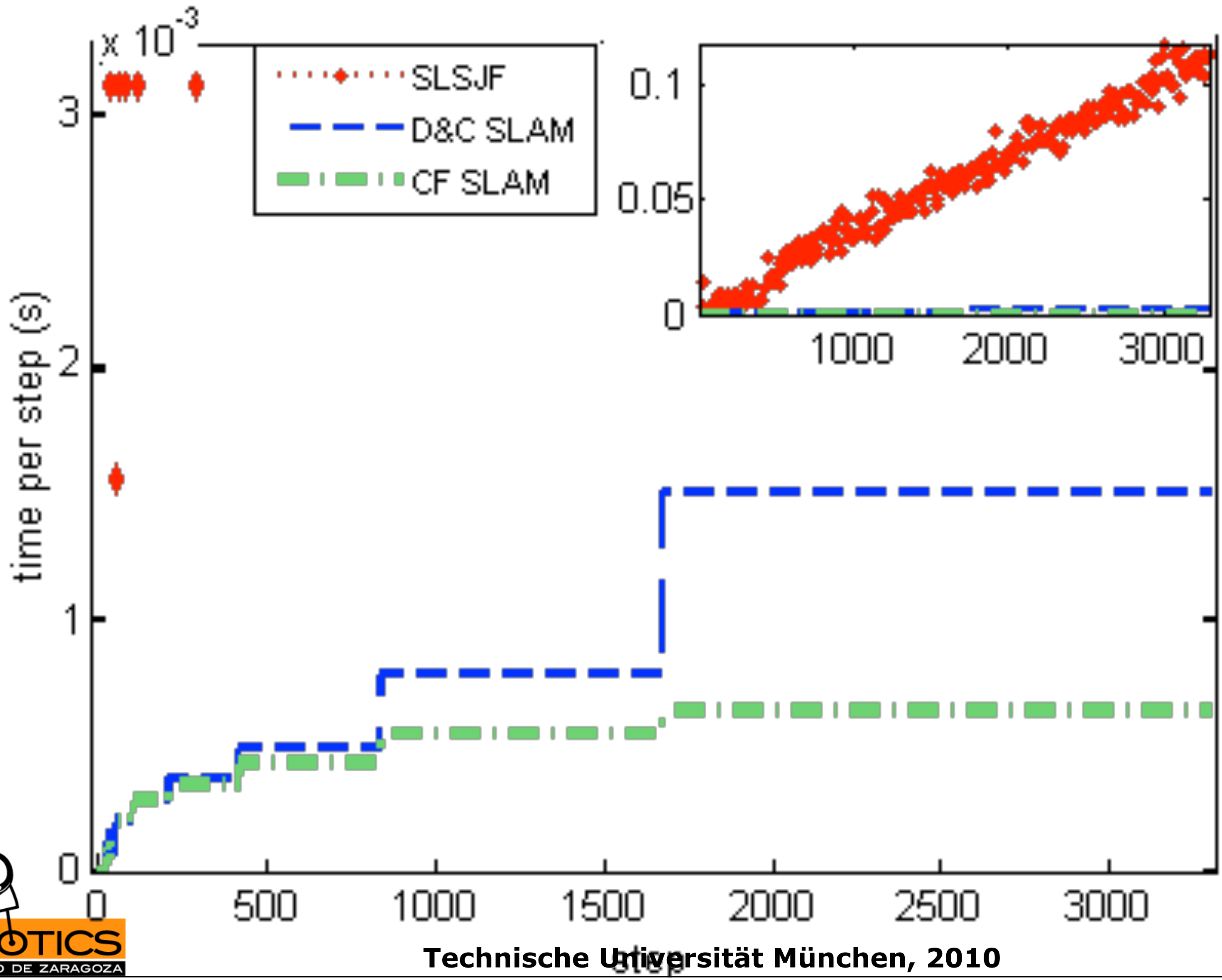
Combined Filter (CF) SLAM

- EKF and EIF Combined

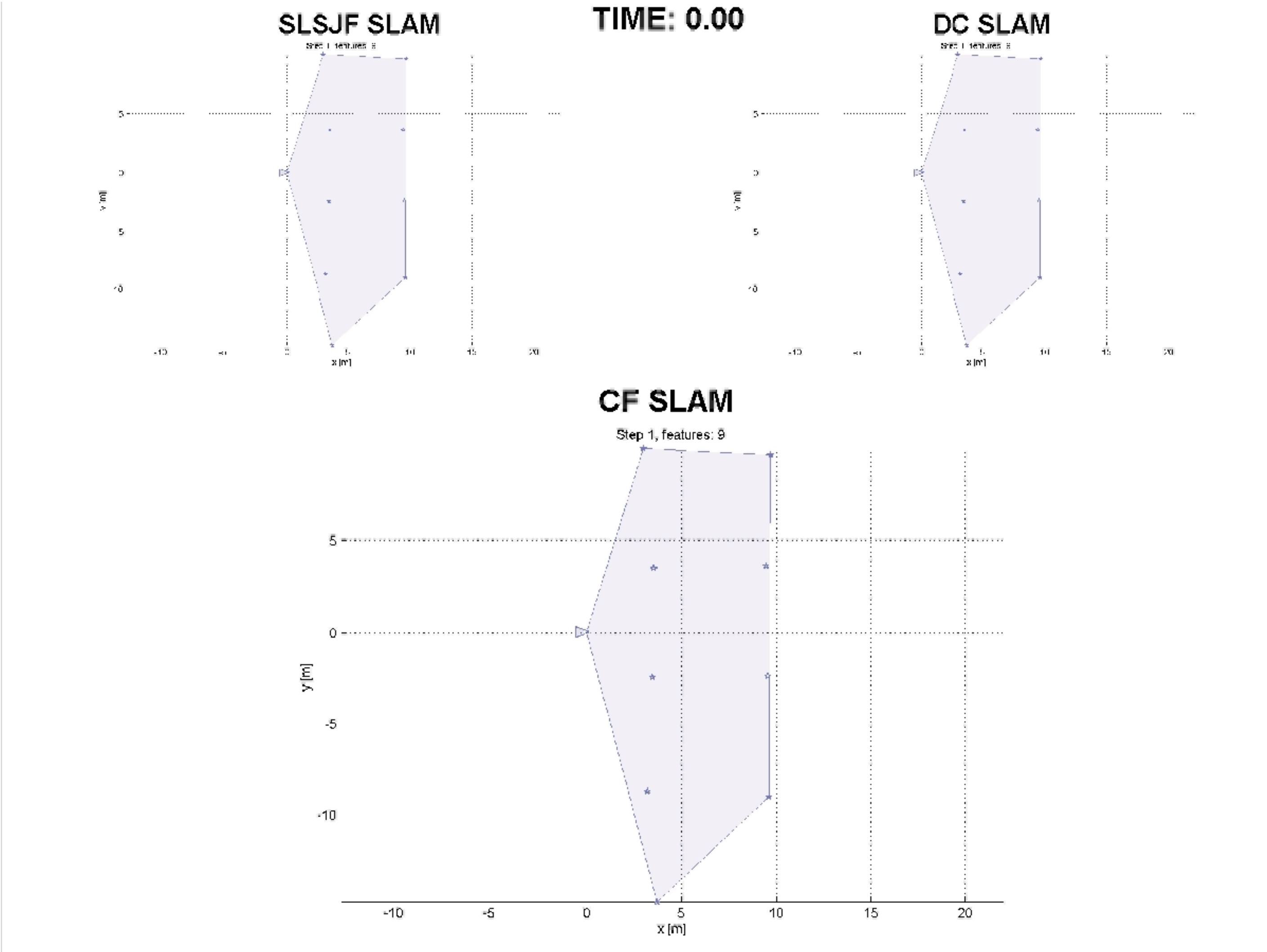


Tec

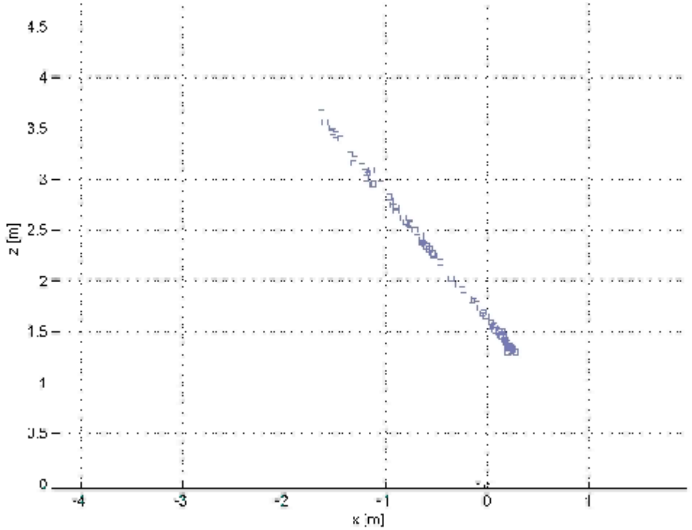
CF SLAM is sublinear

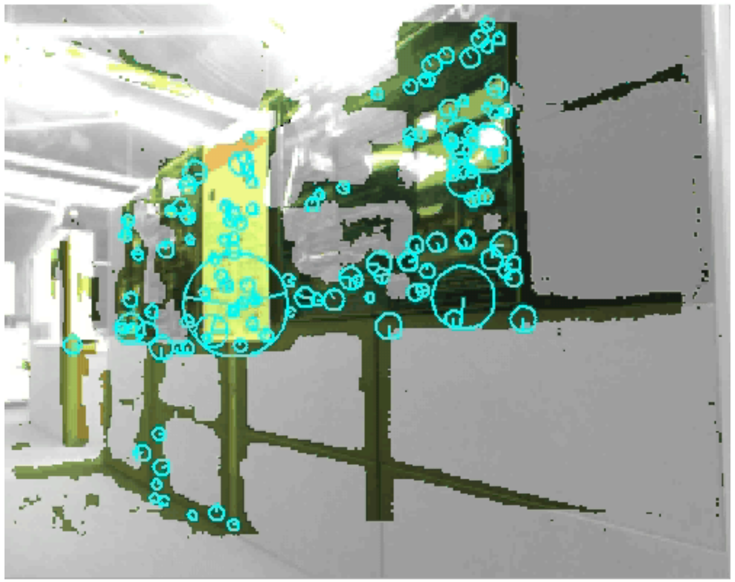


Map updates only

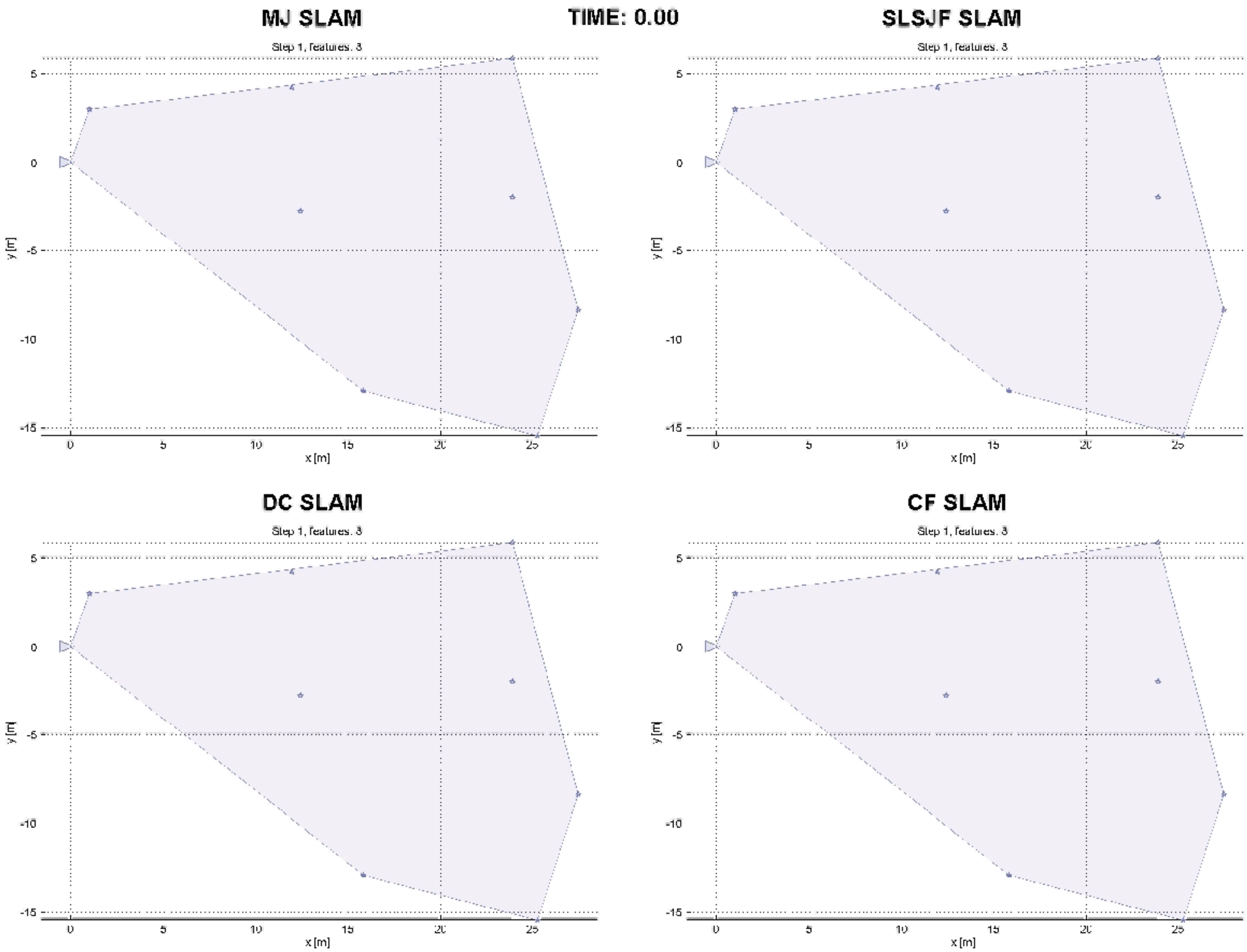


Rose Building, Sydney, with stereo



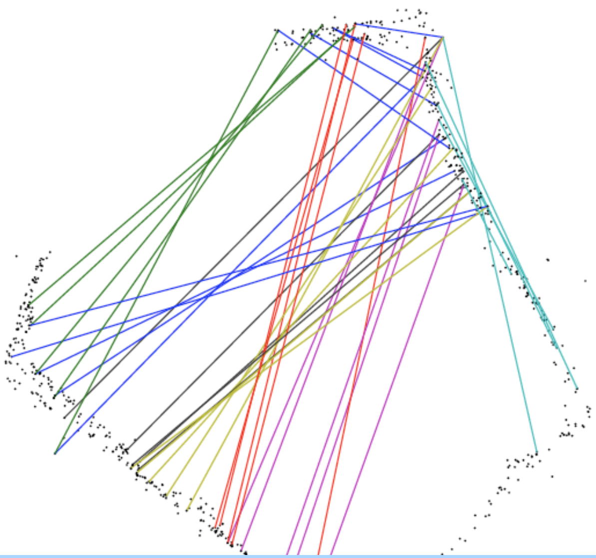


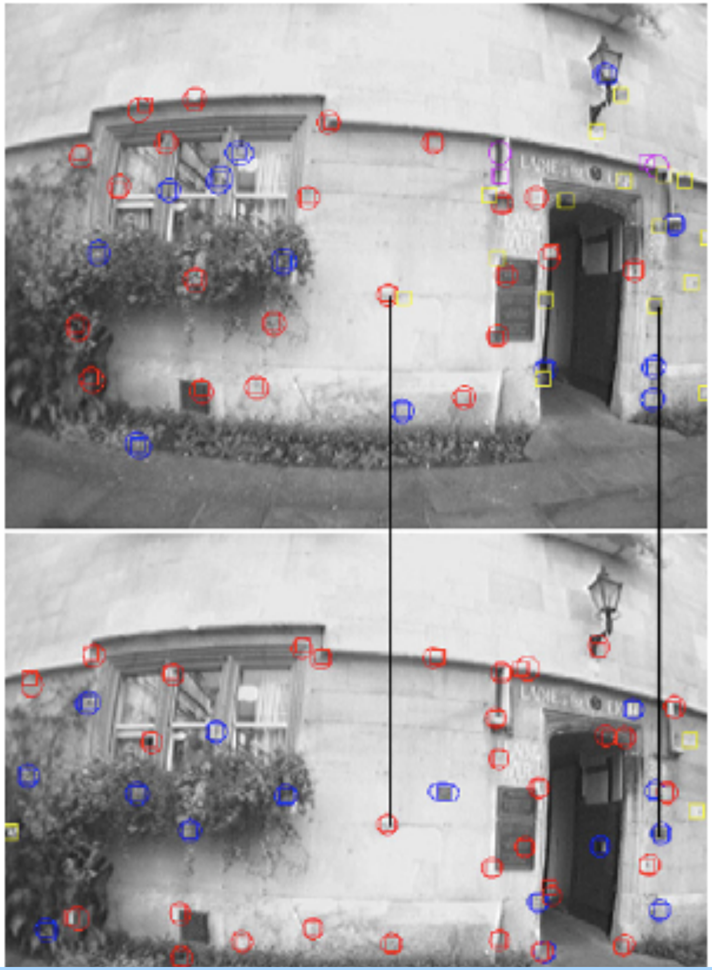
Including data association



The Loop Closing Problem (with B. Williams, Oxford)

- **Map-to-map matching:** Some risk of false negatives
- High risk of false positives with few correspondences

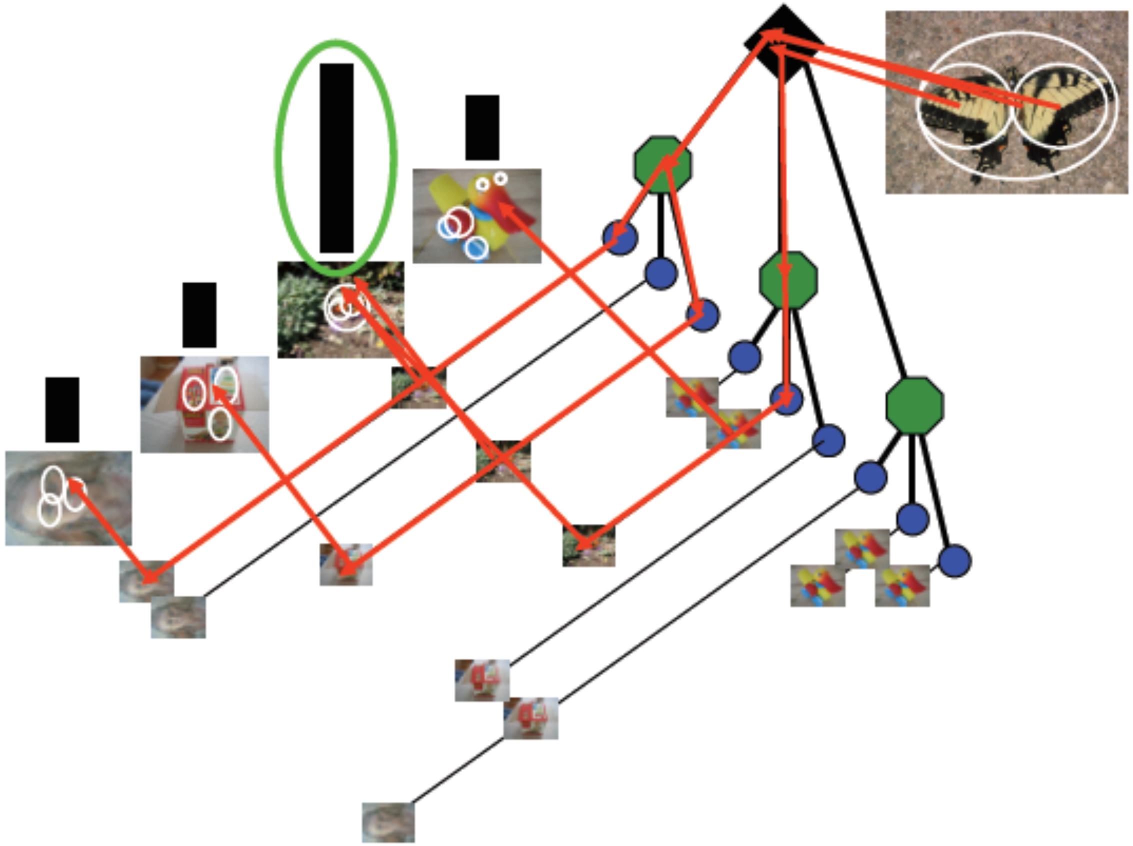




B. Williams, M. Cummins, J. Neira, P. Newman, I. Reid, J. D. Tardos
A comparison of loop closing techniques in monocular SLAM
Robotics and Autonomous Systems, December 2009.

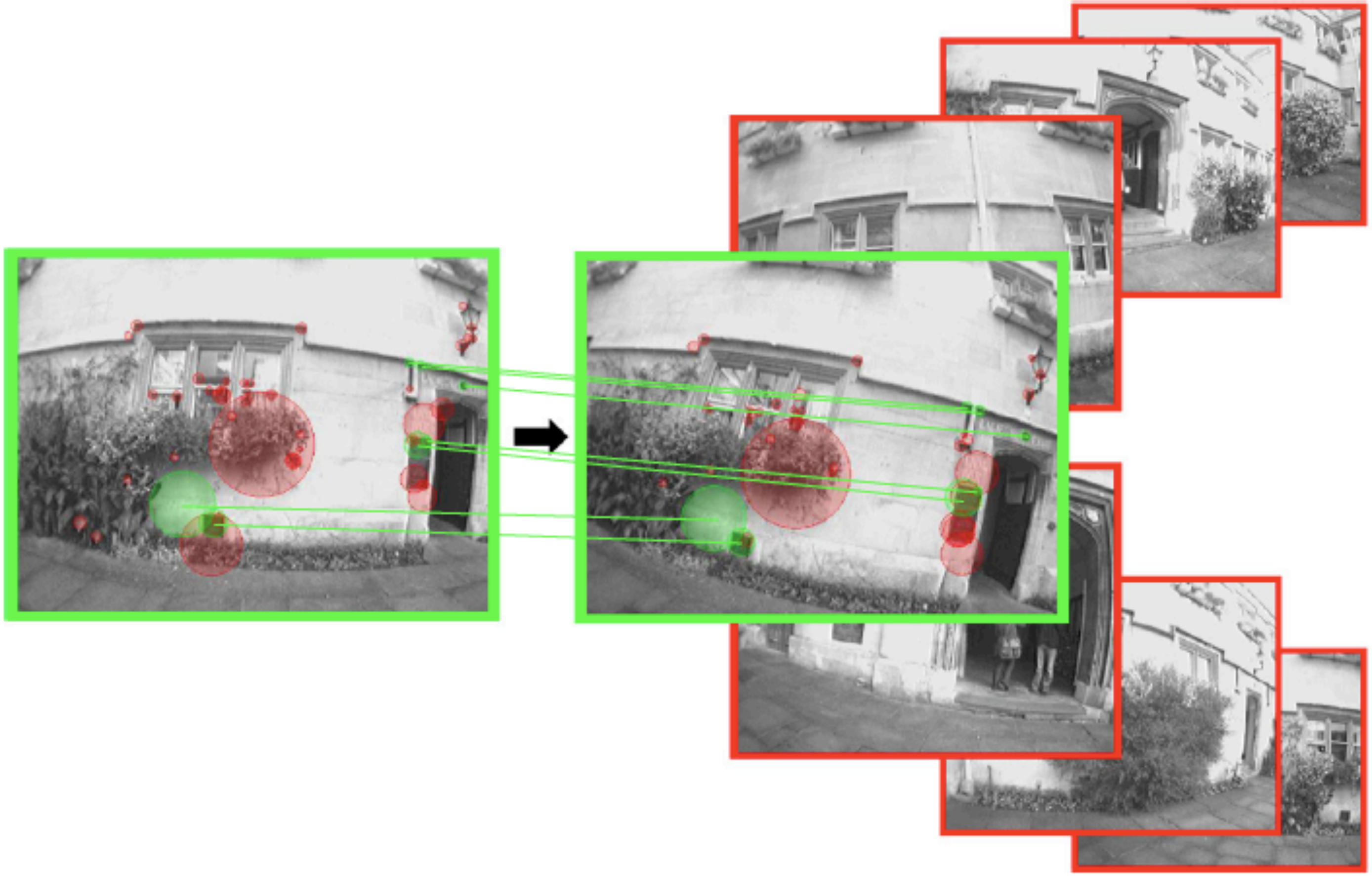
Bags-of-words

- Fast image registration against a database of previous images



Bags-of-words

- Fast image registration against a database of previous images



Bags-of-words

- Some risk of false positives

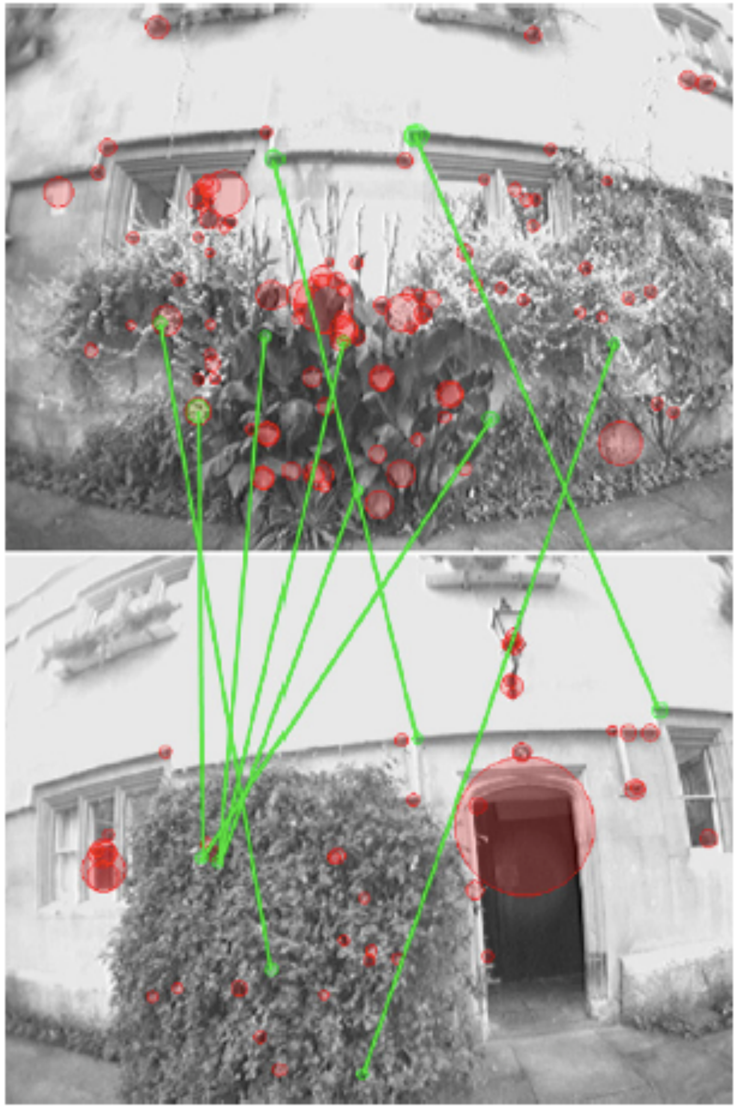
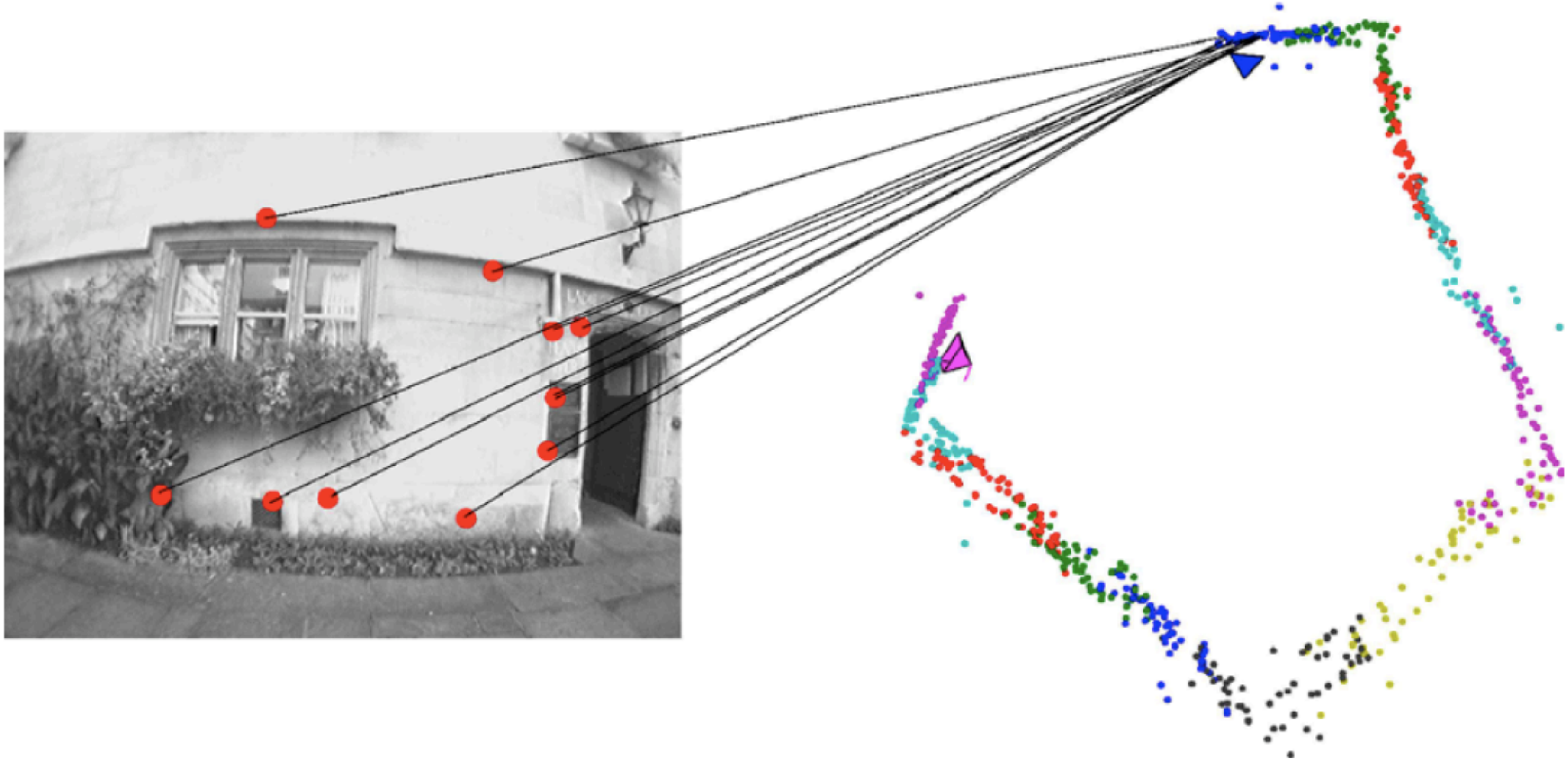
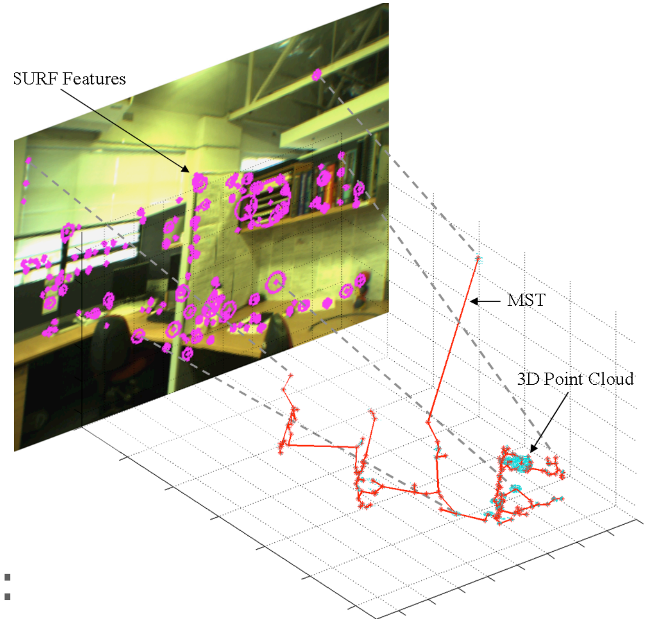


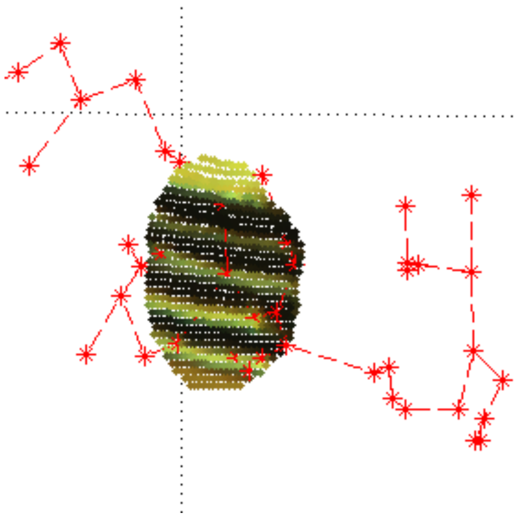
Image-to-Map

- No false positives



Data association: Conditional Random Fields





- Local:
  - SURF description vector
  - Local shape (PCA)
  - Curvature (SVD)
  - Geodesic: sum of Euclean distances to neighbors (or order 1, 2 and 3)
- Relative (pairwise):
  - Euclidean Distance

The RAWSEEDS project

- European FP6 project
  - University of Freiburg
  - Politenico di Milano
  - Universitá degli Studi di Milano
  - Universidad de Zaragoza

- Aim: to build benchmarking tools for robotic systems.

- Publication of a comprehensive, high-quality Benchmarking Toolkit composed of:
  - high-quality multisensor datasets, with associated ground truth;
  - Benchmark Problems based on the datasets;
  - Benchmark Solutions for the problems

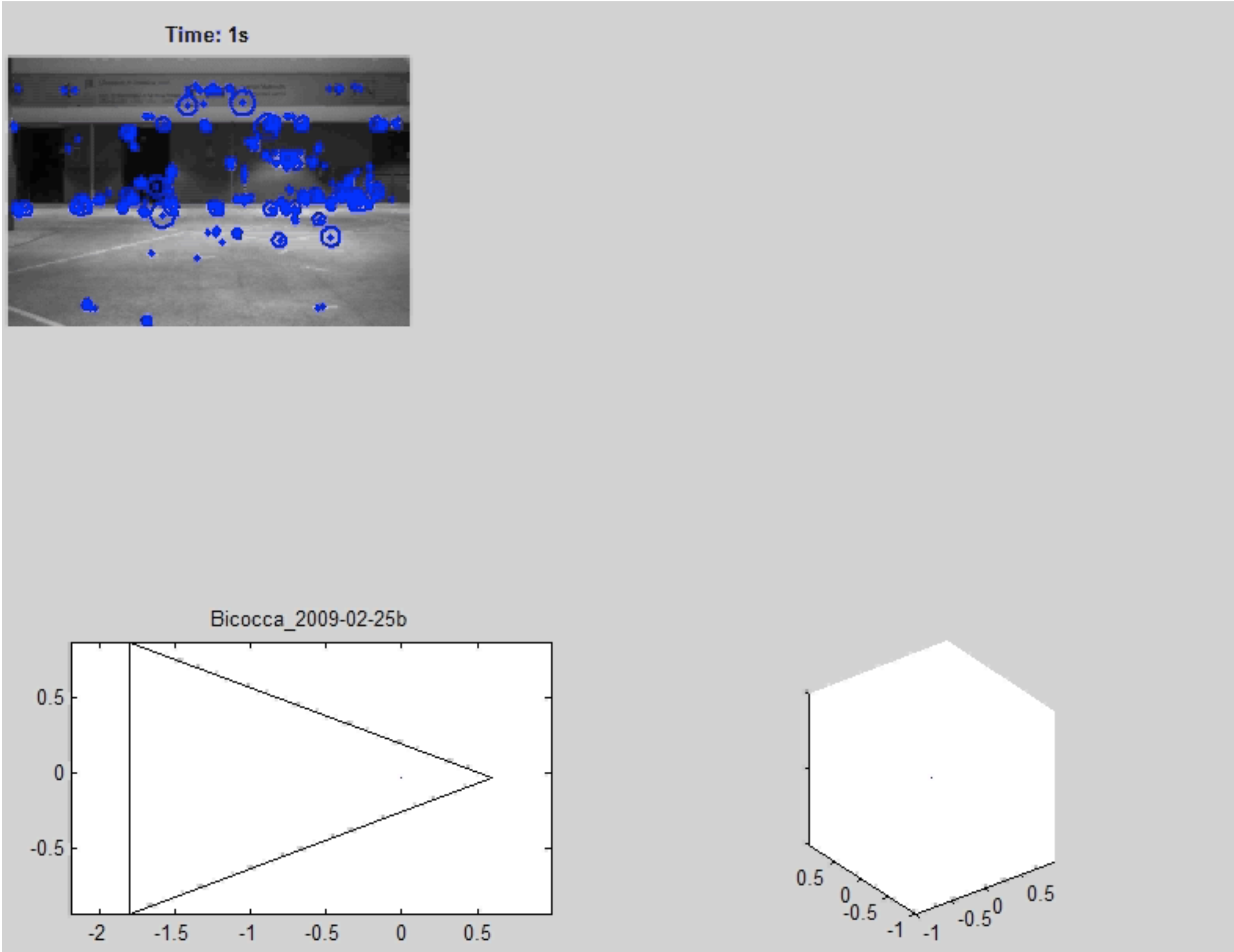
RAWSEEDS: indoor

Current Candidate Accepted

BoW

CRF Matching

30x



Experiments - Mixed

Current

Candidate

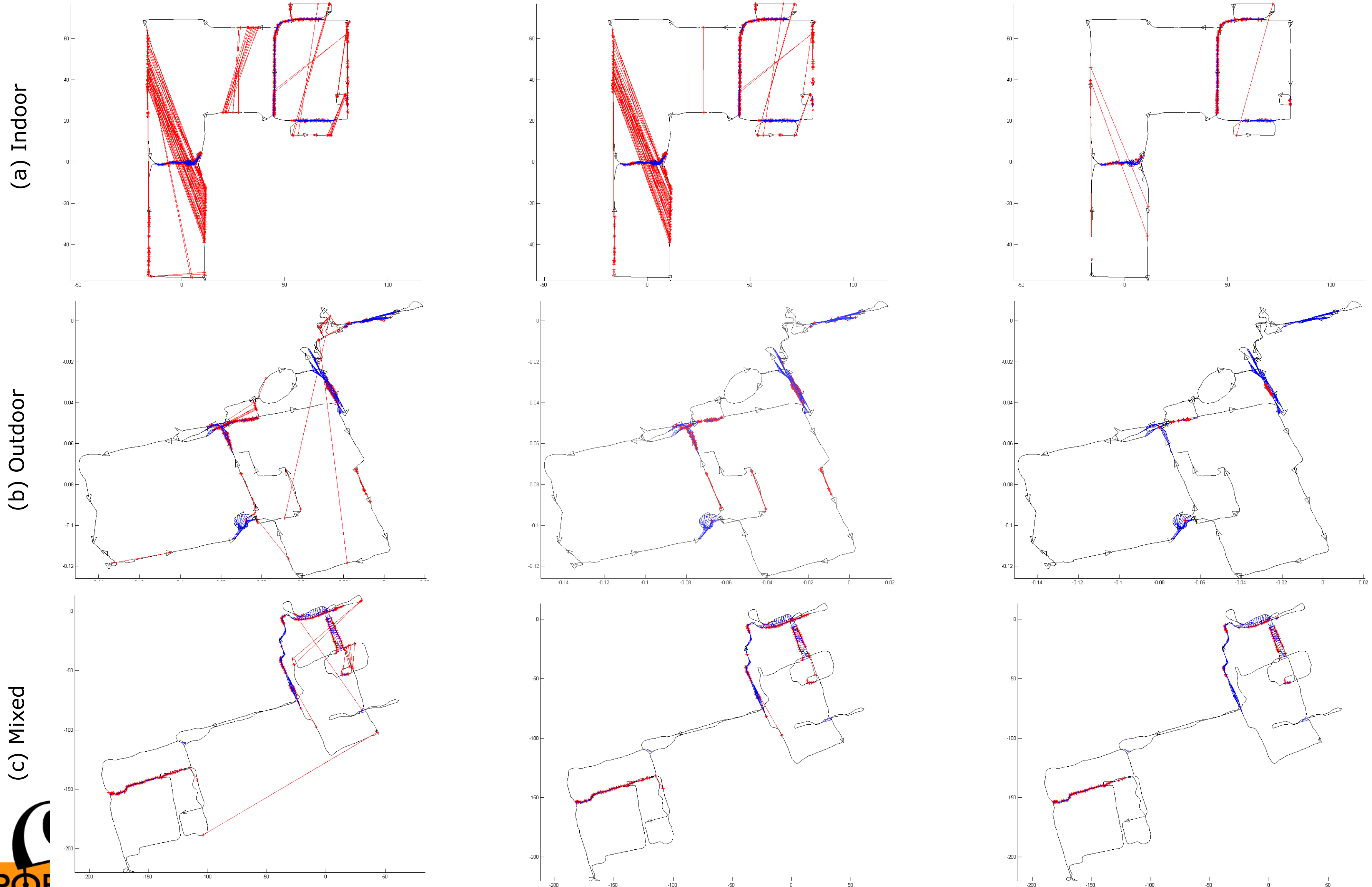
Accepted

BoW

CRF Matching

30x

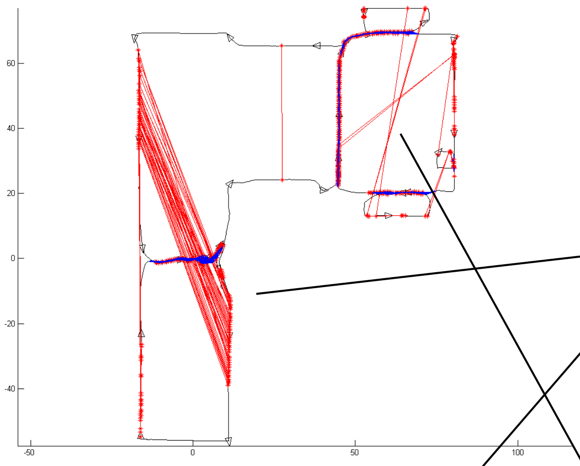
BoW | BoW+epipolar | BoW+CRF

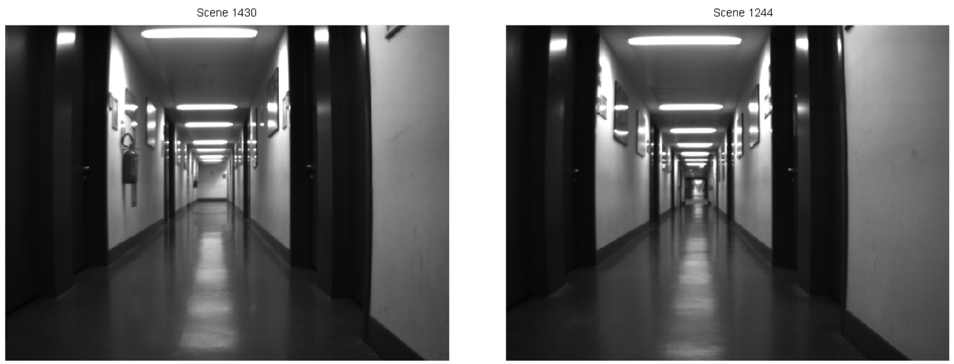


False positives

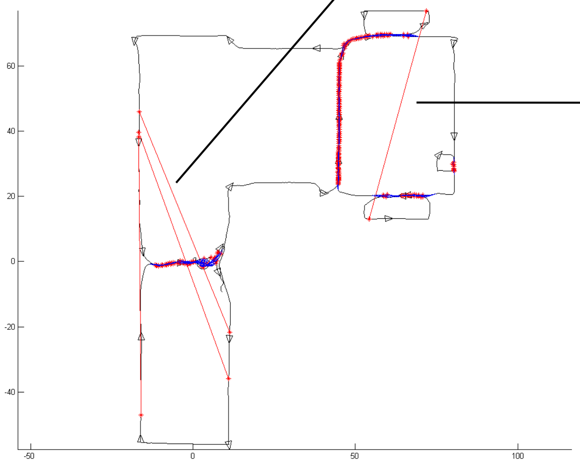
(a) Indoor

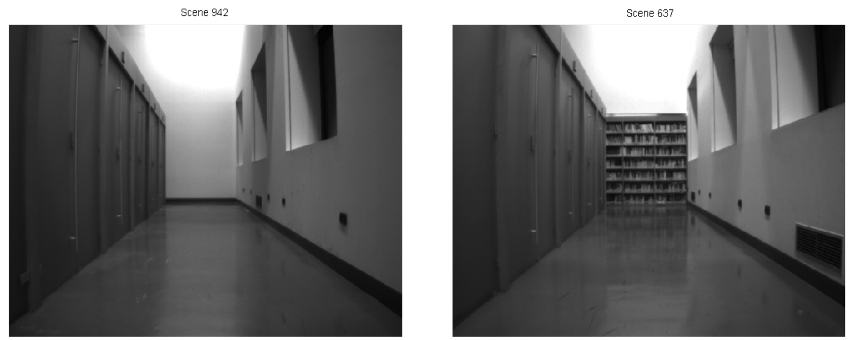
BoW+epipolar

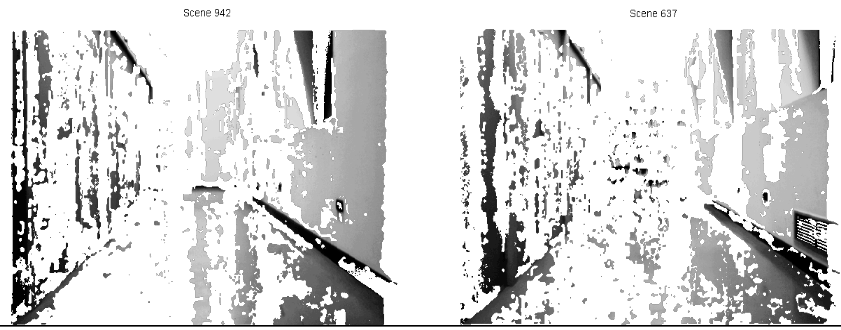




BoW+CRF





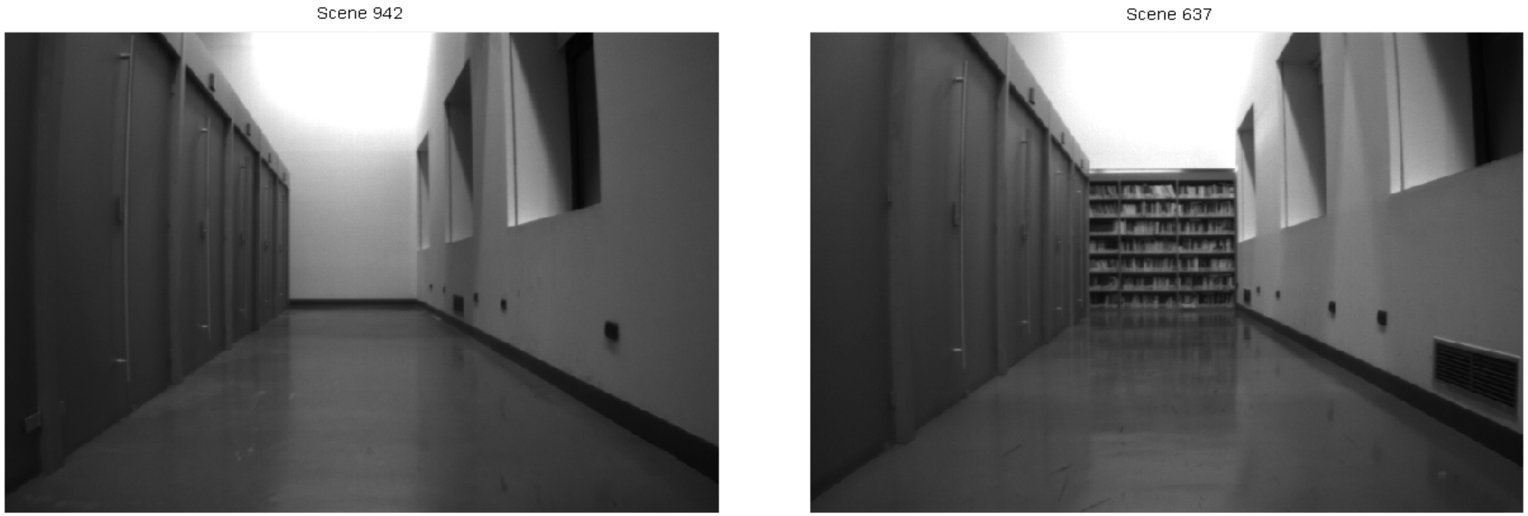


Bow+CRFs: False positives

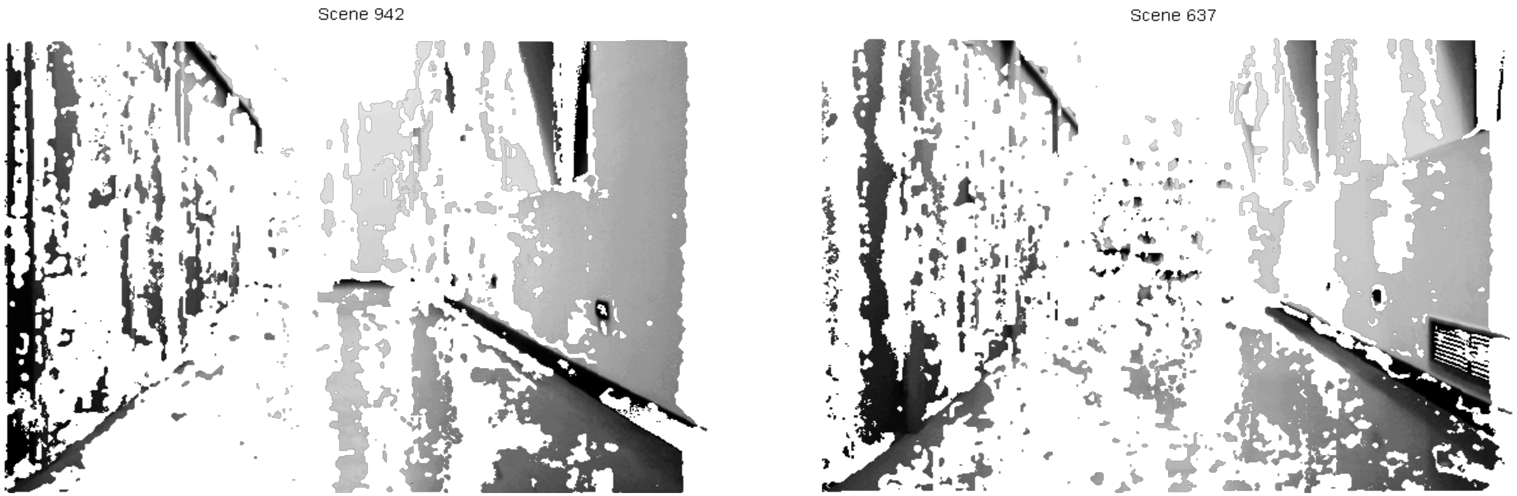
Scene 1430

Scene 1244

Bow+CRFs: False positives



(b) Library



Bow+CRFs: False positives

Scene 292

Scene 219

Bow+Epipolar: False positives

Scene 1317

Scene 1126

Conclusions

- Estimation methods are well understood:
  - EKF, EIF, SAM, TJTFs, graphSLAM, bundle adjustment

H. Strasdat and J. M. M. Montiel, A. Davison
Real-Time Monocular SLAM: Why Filter?

Best Vision Paper at ICRA 2010.

  - *"while filtering may have a niche in systems with low processing resources, in most modern applications keyframe optimisation gives the most accuracy per unit of computing time."*

ROBOTICS
UNIVERSIDAD DE ZARAGOZA

Is SLAM solved?

- Interview by Udo Frese (U. Bremen) with S. Thrun, J. Neira, to appear in Journal **Künstliche Intelligenz**:

- Maybe for indoor static environments, but...

- **SLAM is NOT solved for:**
  - Dynamic SLAM
  - Semantic SLAM
  - Lifelong SLAM

- Data association is still a challenging problem