Environment Modelling for Robots using Only Cameras

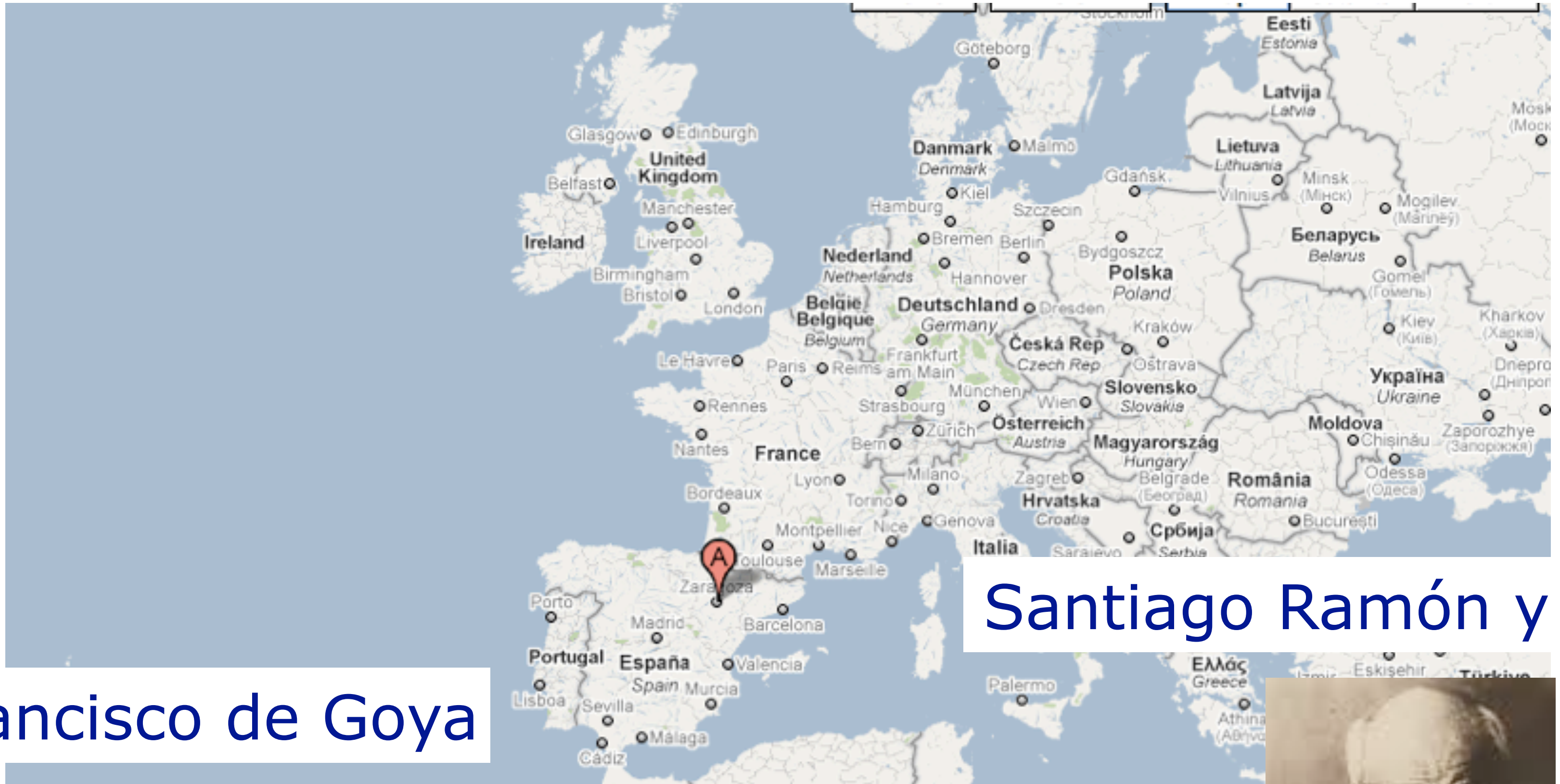
José Neira

Universidad de Zaragoza

Joint work with:

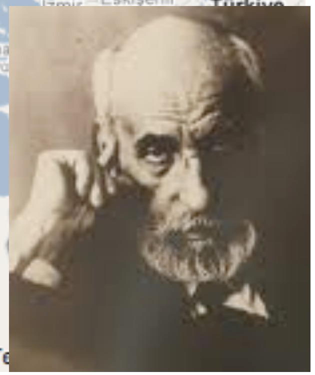
- **Universidad de Zaragoza**
  - César Cadena, Lina Paz, Pedro Piníes, Juan Tardós
- **University of Oxford**
  - Brian Williams, Paul Newman, Ian Reid
- **Imperial College London**
  - Andrew Davison
- **Massachusetts Institute of Technology**
  - Michael Kaess, John Leonard
- **National University of Ireland**
  - John McDonald

Zaragoza, where is that?



Santiago Ramón y Cajal

Francisco de Goya



Universidad Zaragoza
1542

Motivation (late 1980s)

Simultaneous Localization and Mapping

Is it possible to use a vehicle, starting at an

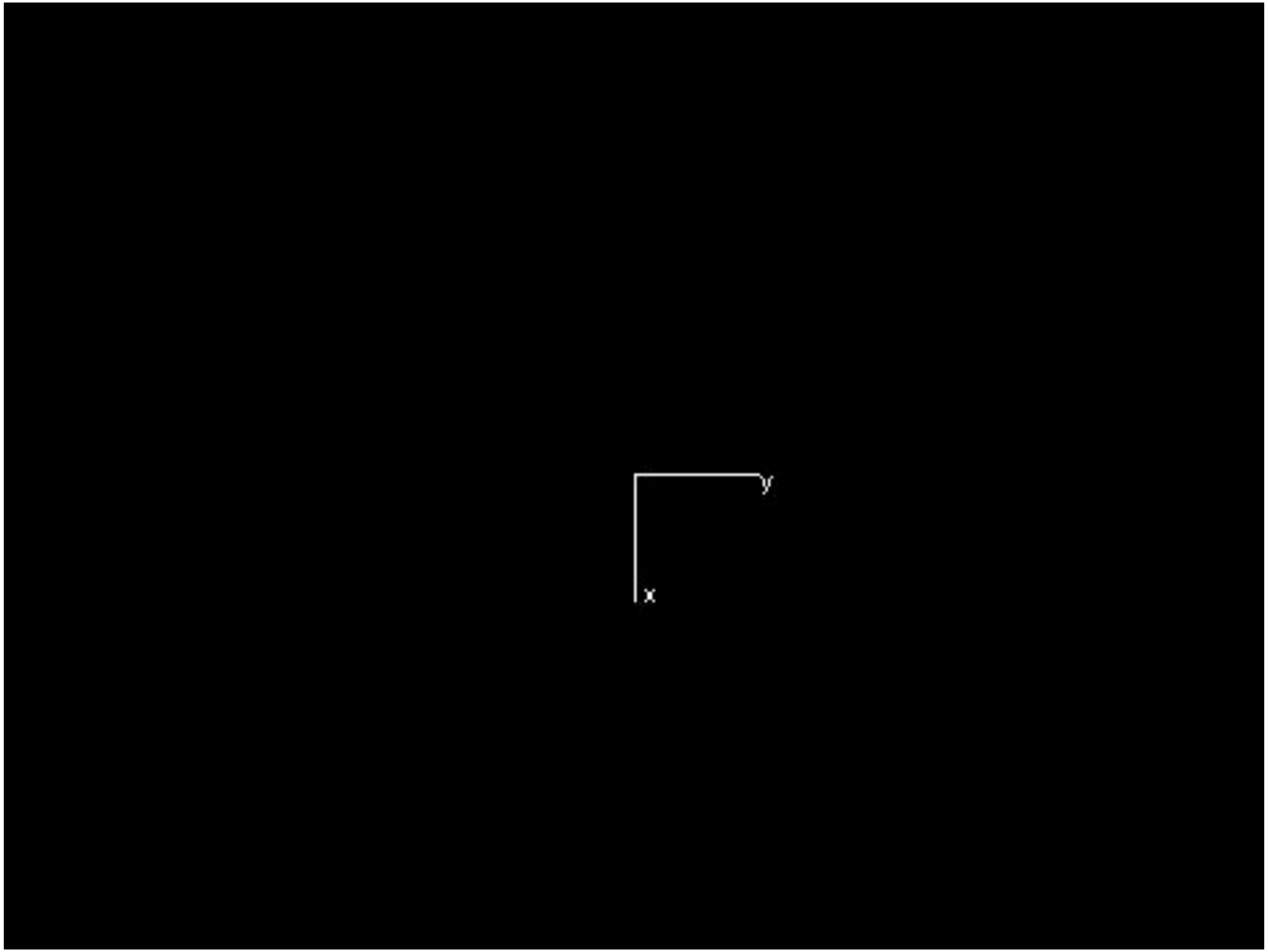
- **unknown initial location**, in an
- **unknown environment**, to
- **incrementally**

build a map of the environment,

- and **at the same time**

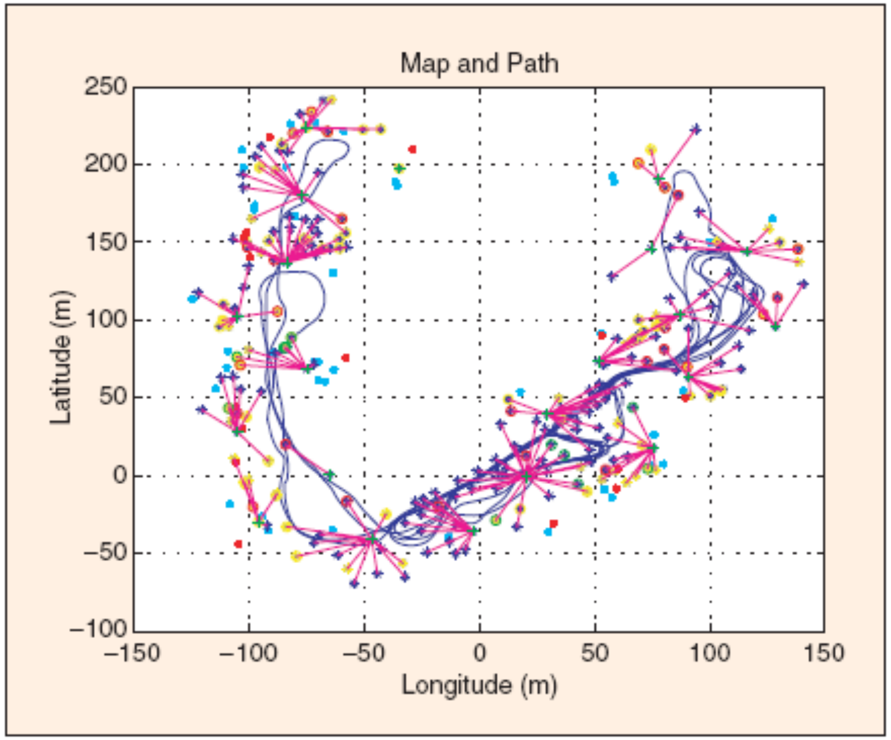
use the map to determine
the vehicle location?

Motivation



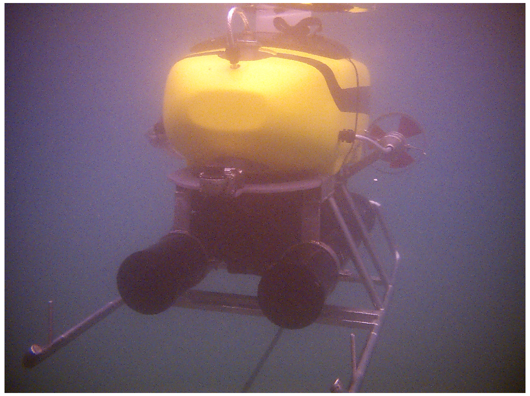
(video: Paul Newman)

Outdoor vehicles



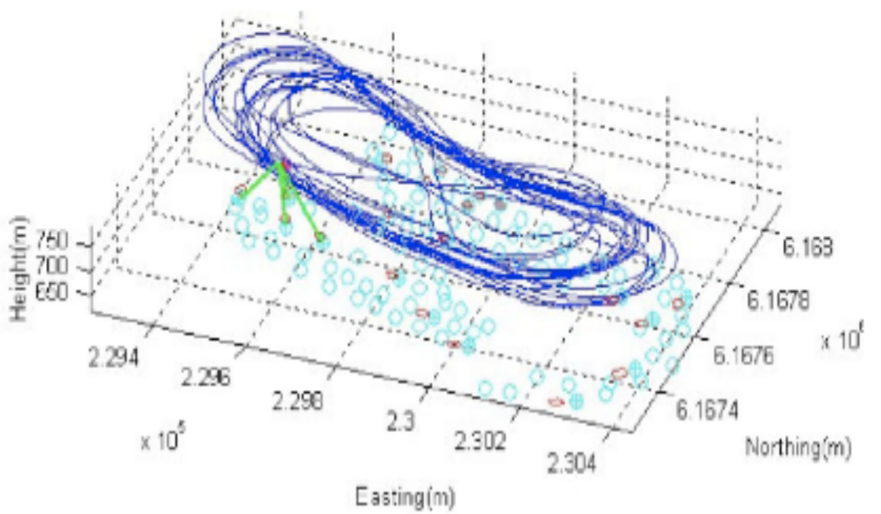
Victoria Park, Univ. Syndey

Underwater, Airborne



Garbi, Univ. Girona, Spain

Brumby, Univ. Syndey



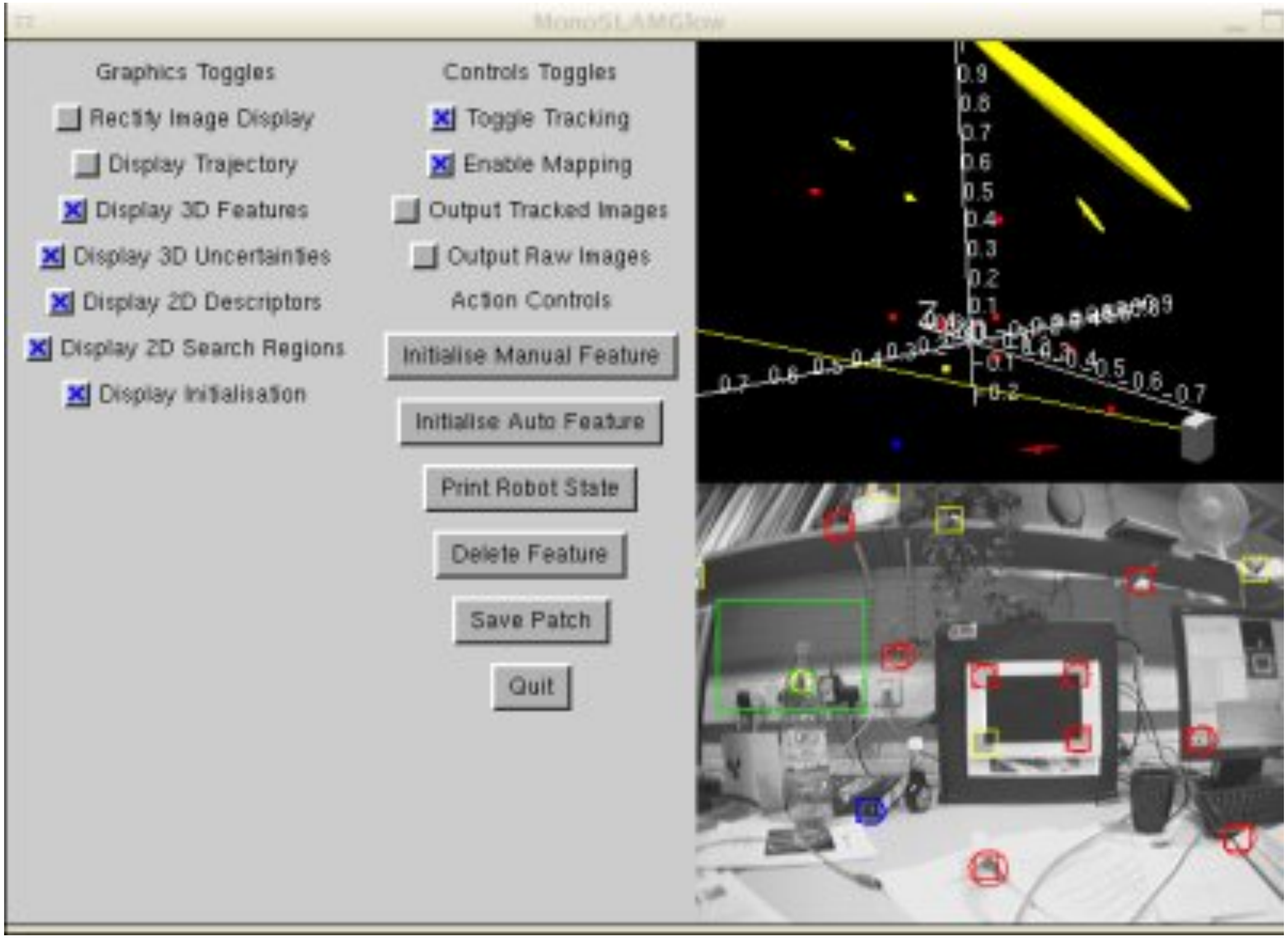
Fundamental issues

- **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
©2007 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 feature 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

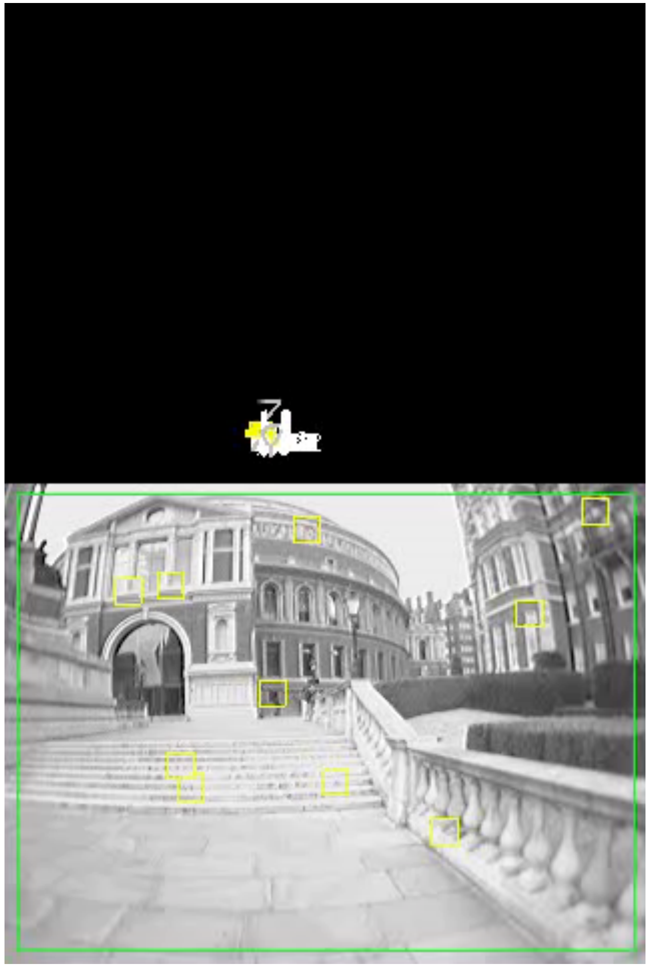
\- Elevation

\- Inverse depth

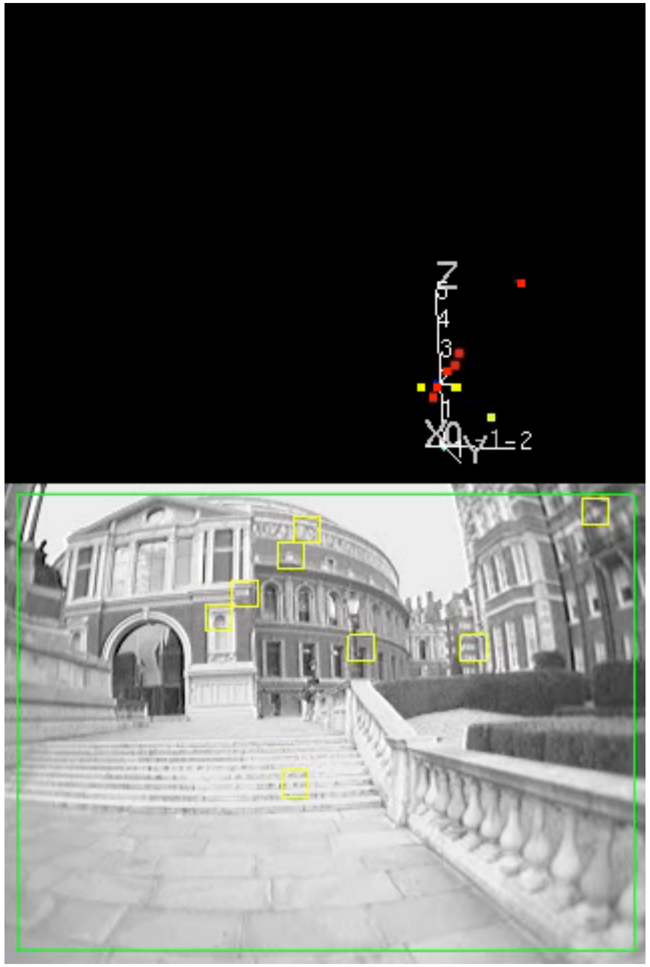
$$\mathbf{y}_i = \begin{pmatrix} x_i \\ y_i \\ z_i \\ \theta_i \\ \phi_i \\ \rho_i \end{pmatrix}$$

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



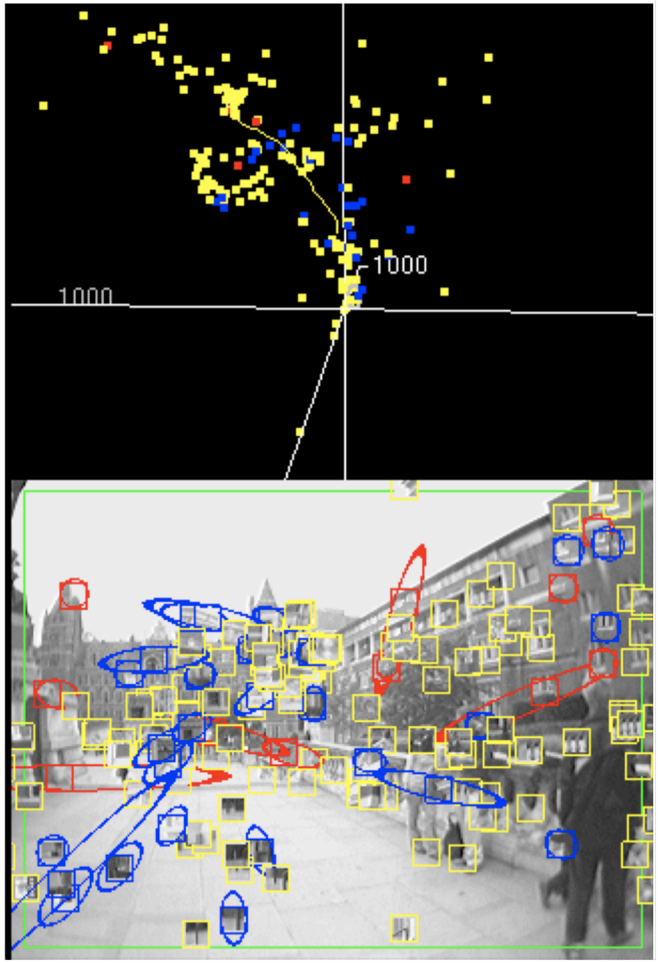
Individual tracks



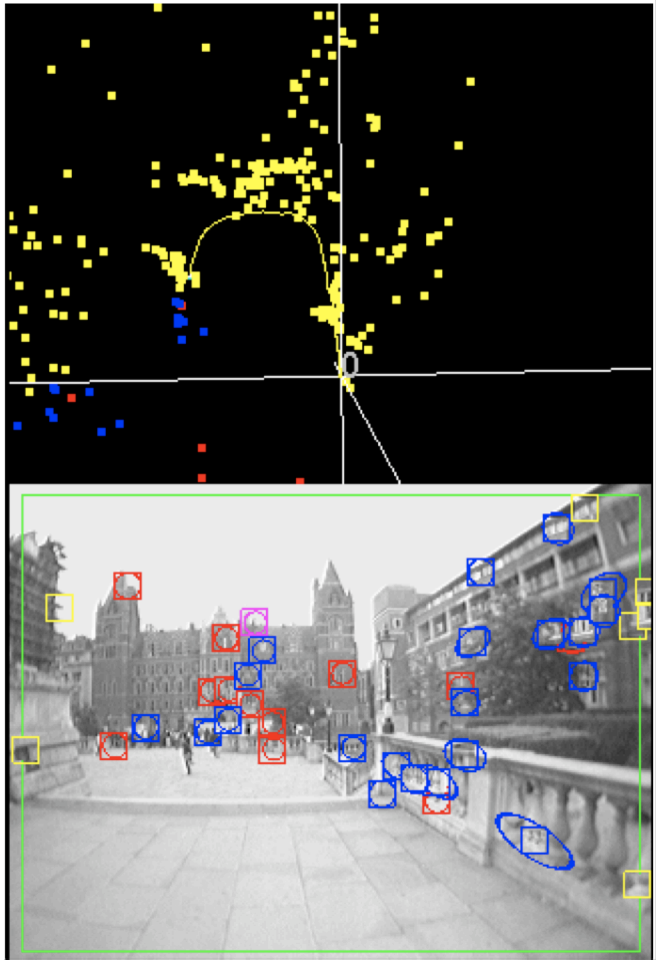
Jointly compatible tracks

Universidad Zaragoza 1542

Nearest neighbor .vs. Joint Compatibility



Individual Tracks

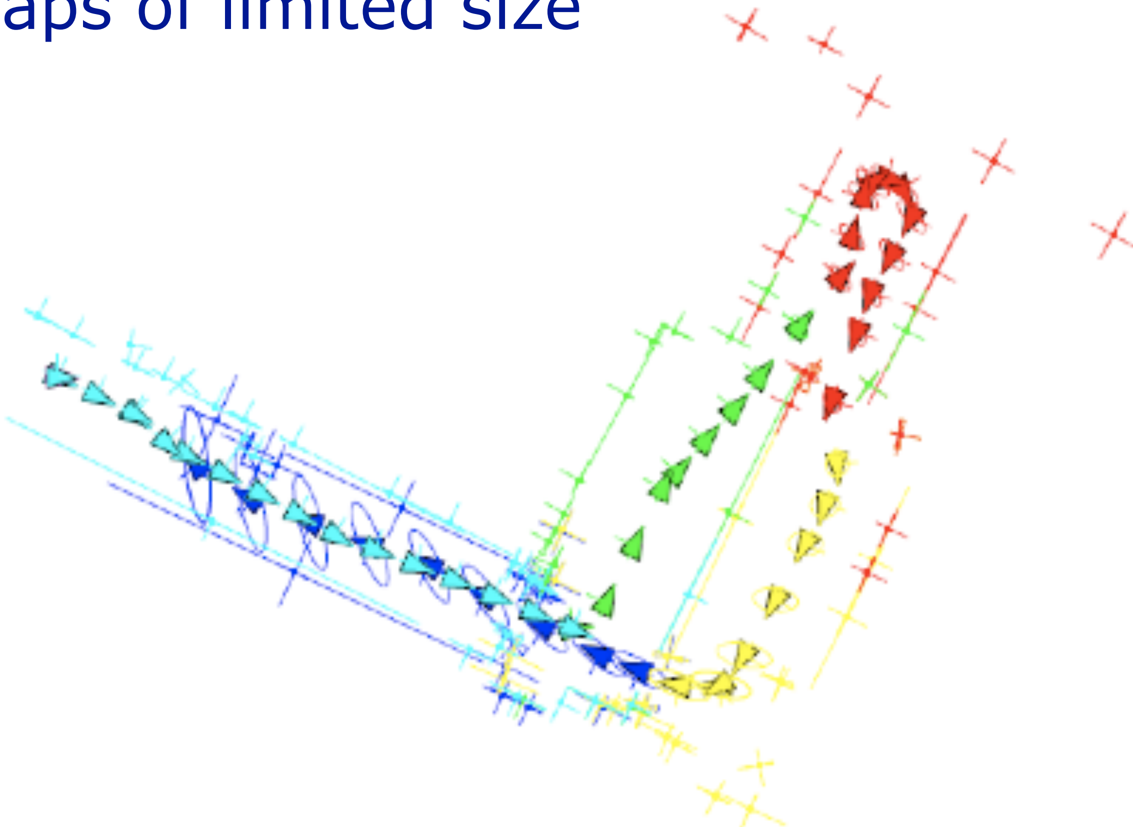


Jointly Compatible Tracks (cost: 2ms)

Universidad Zaragoza 1542

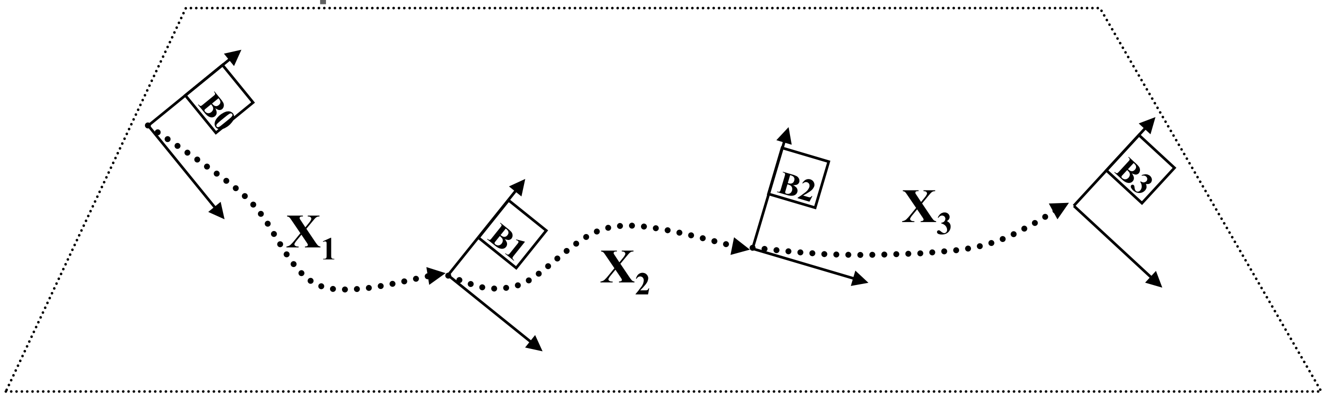
Scalable EKF SLAM:

- Local maps of limited size

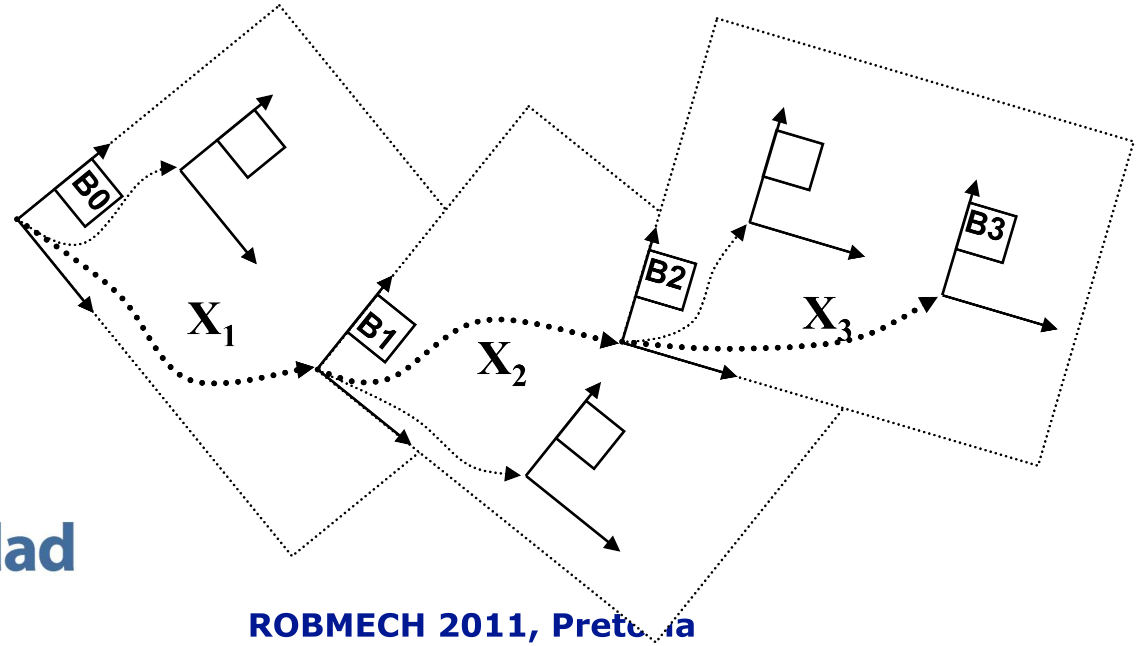


Hierarchical SLAM

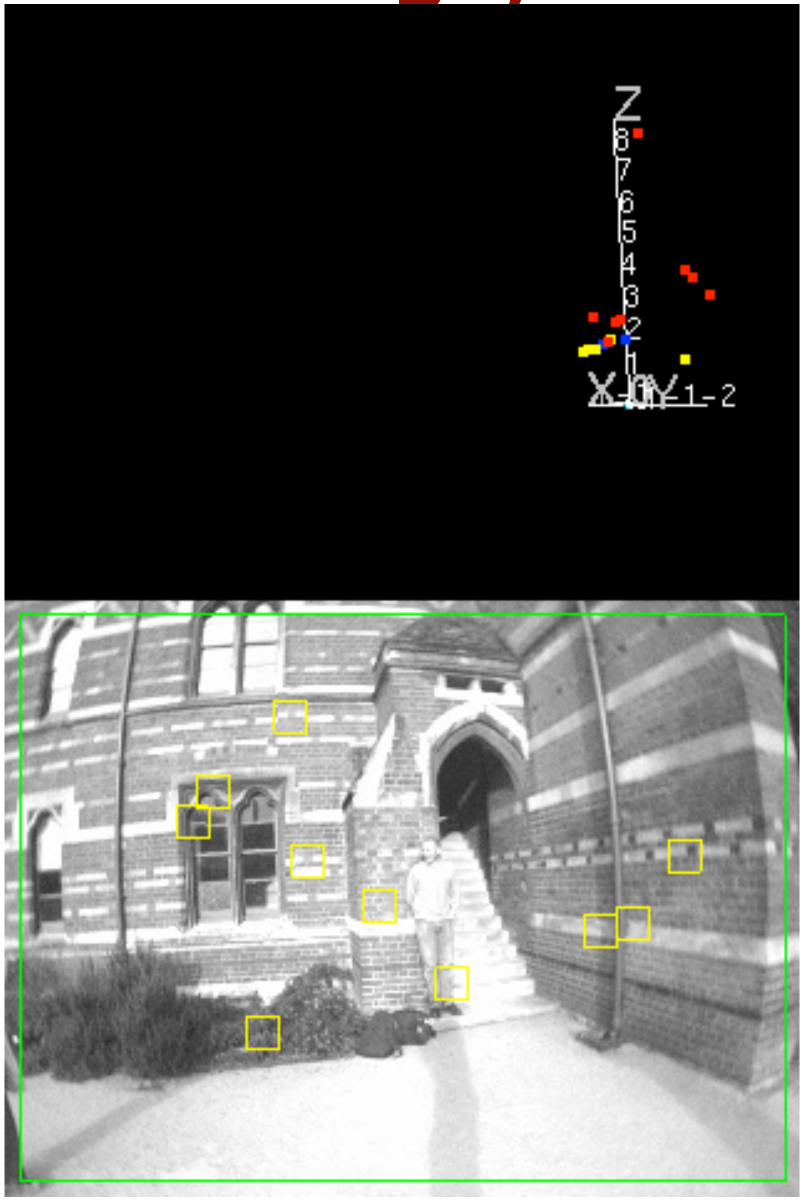
- Global level: adjacency graph and relative stochastic map



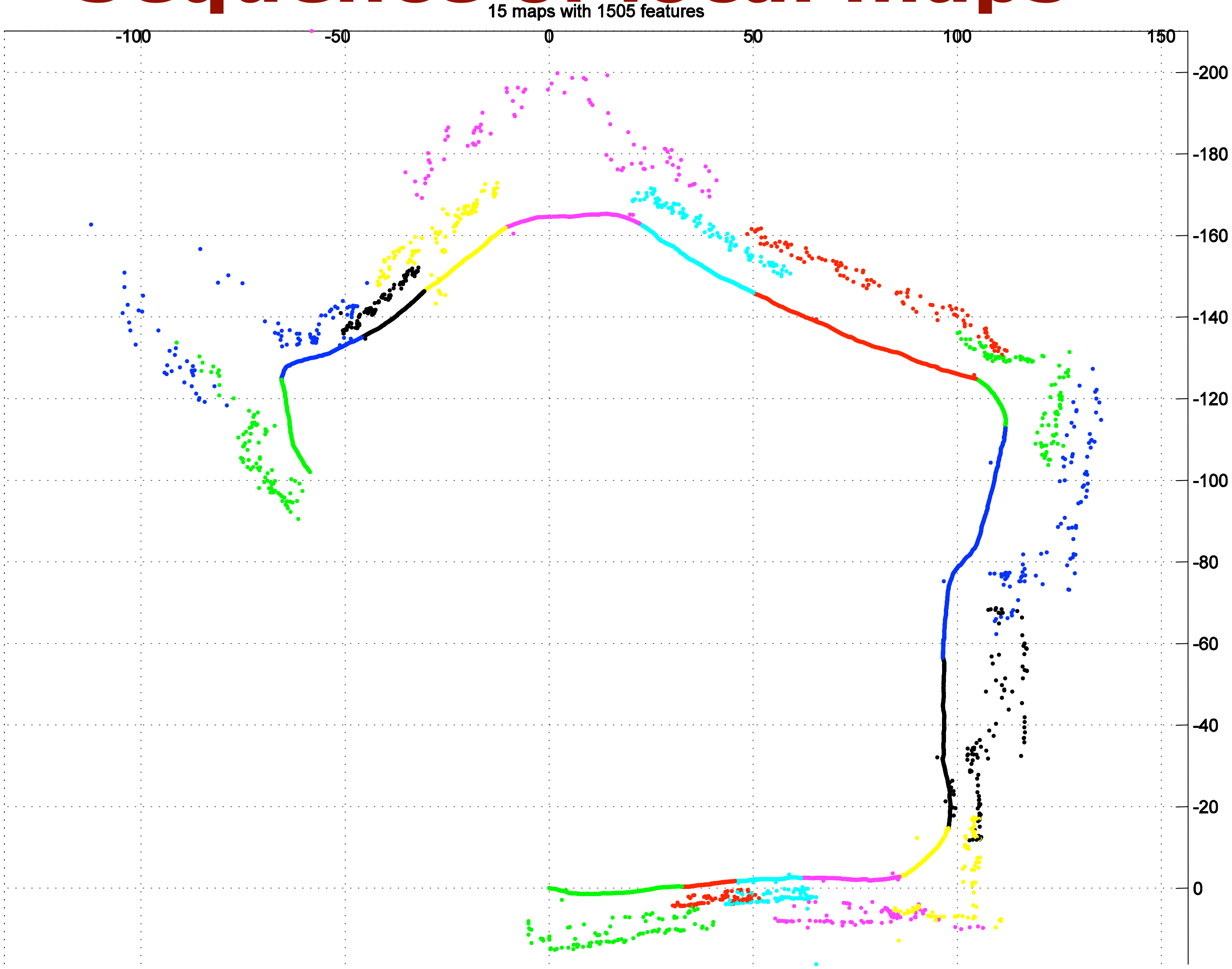
- Local level: statistically independent local maps



Keble College, Oxford

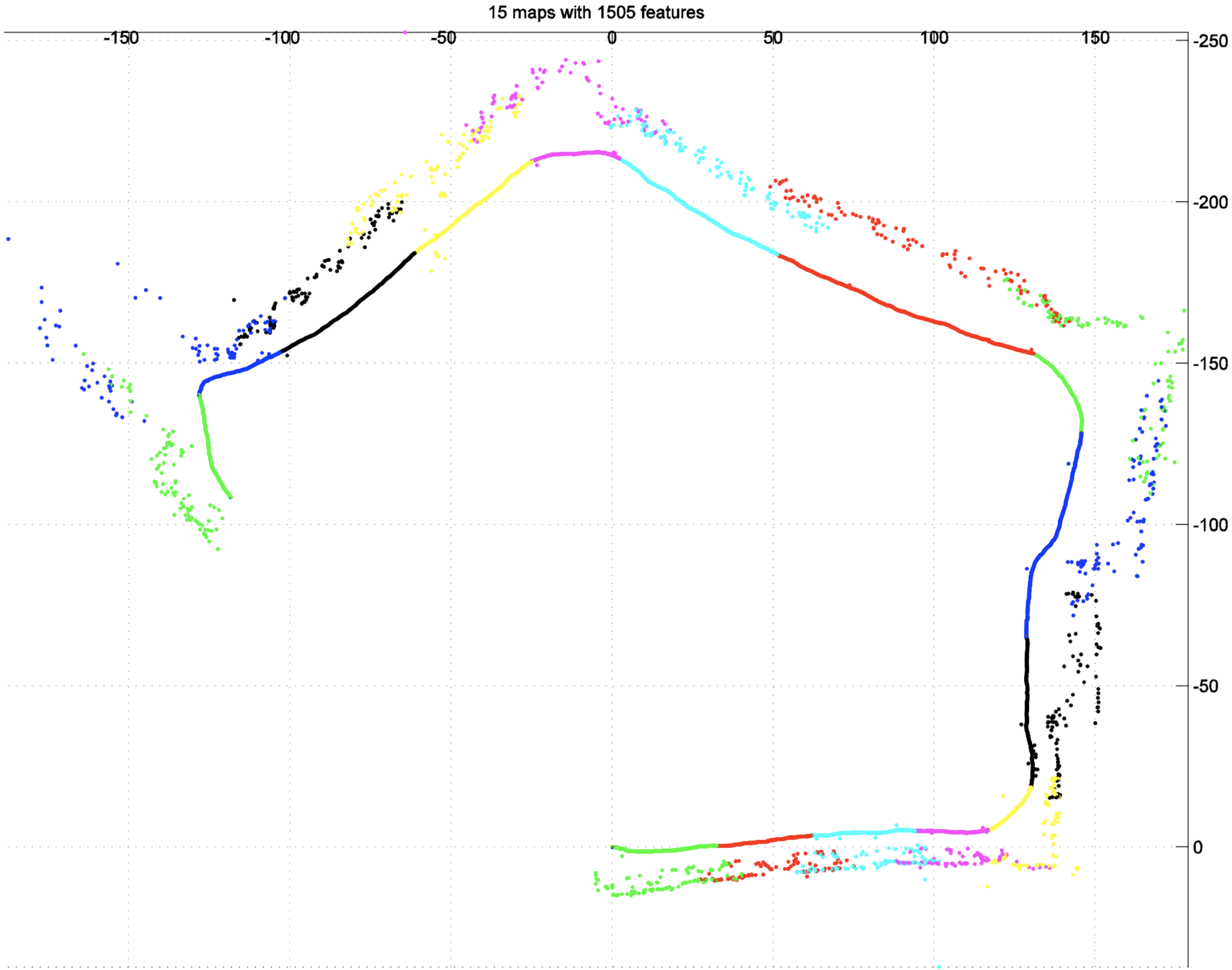


Sequence of local maps

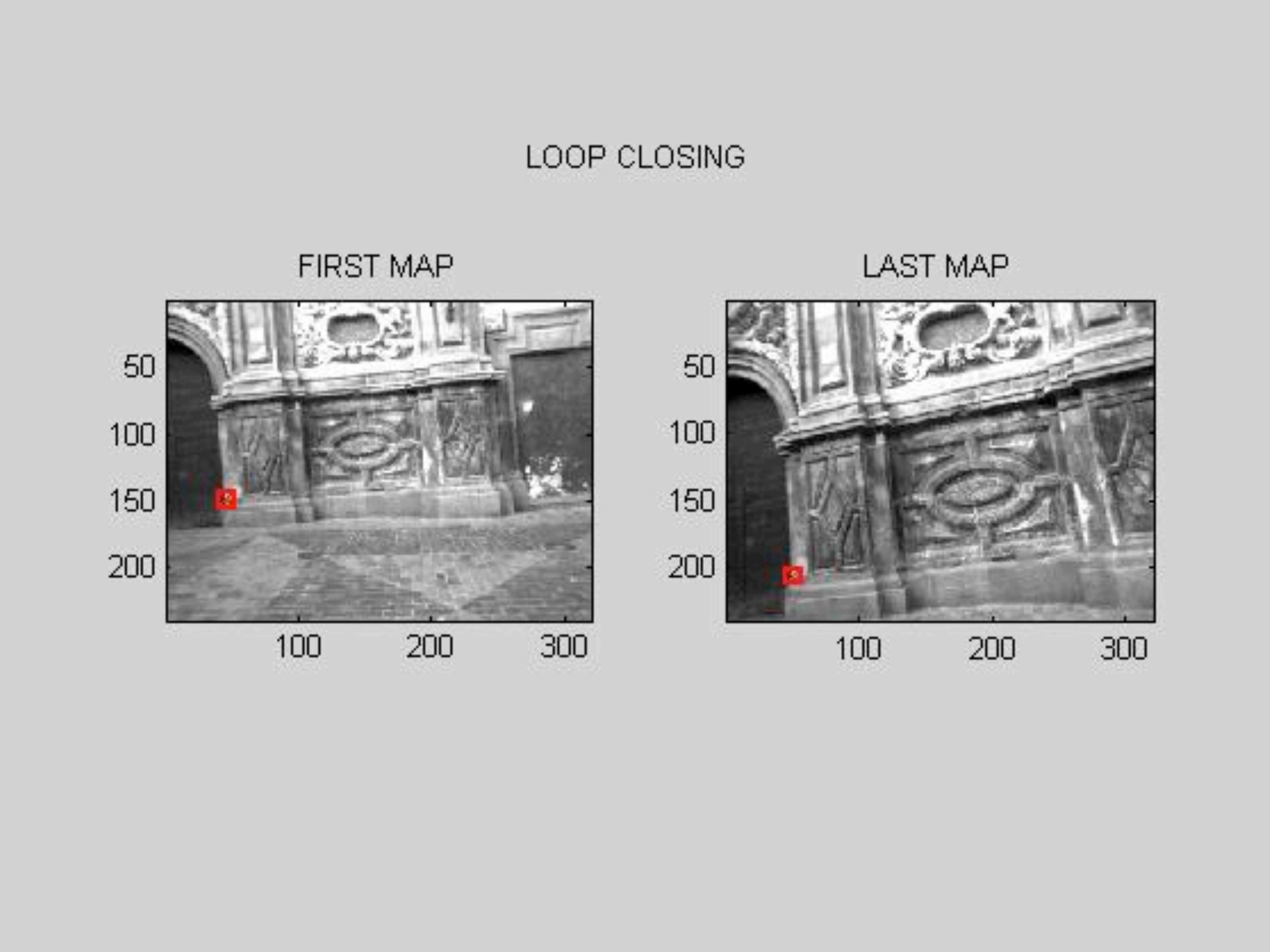


The scale is arbitrary (not observable)

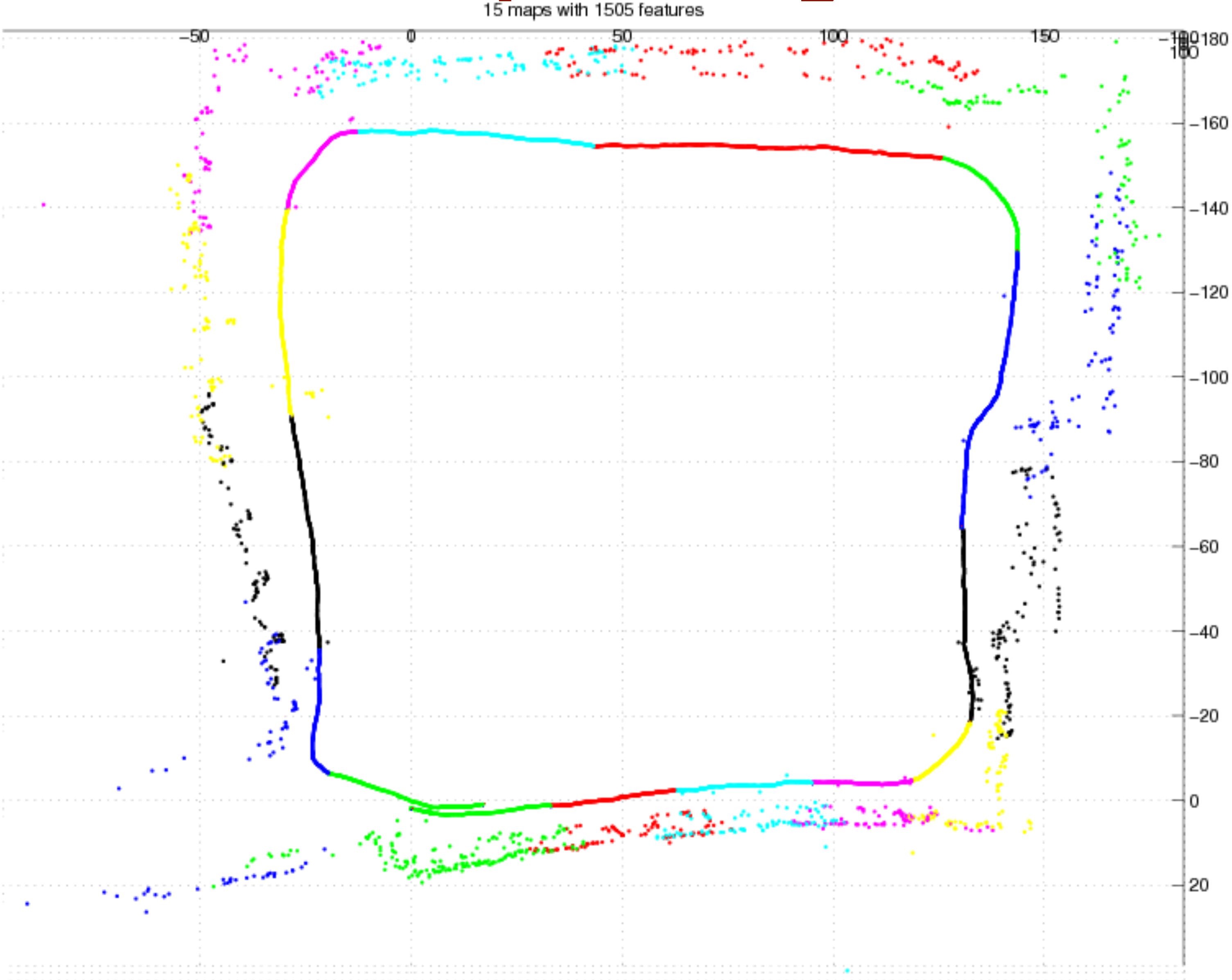
With scale compensation



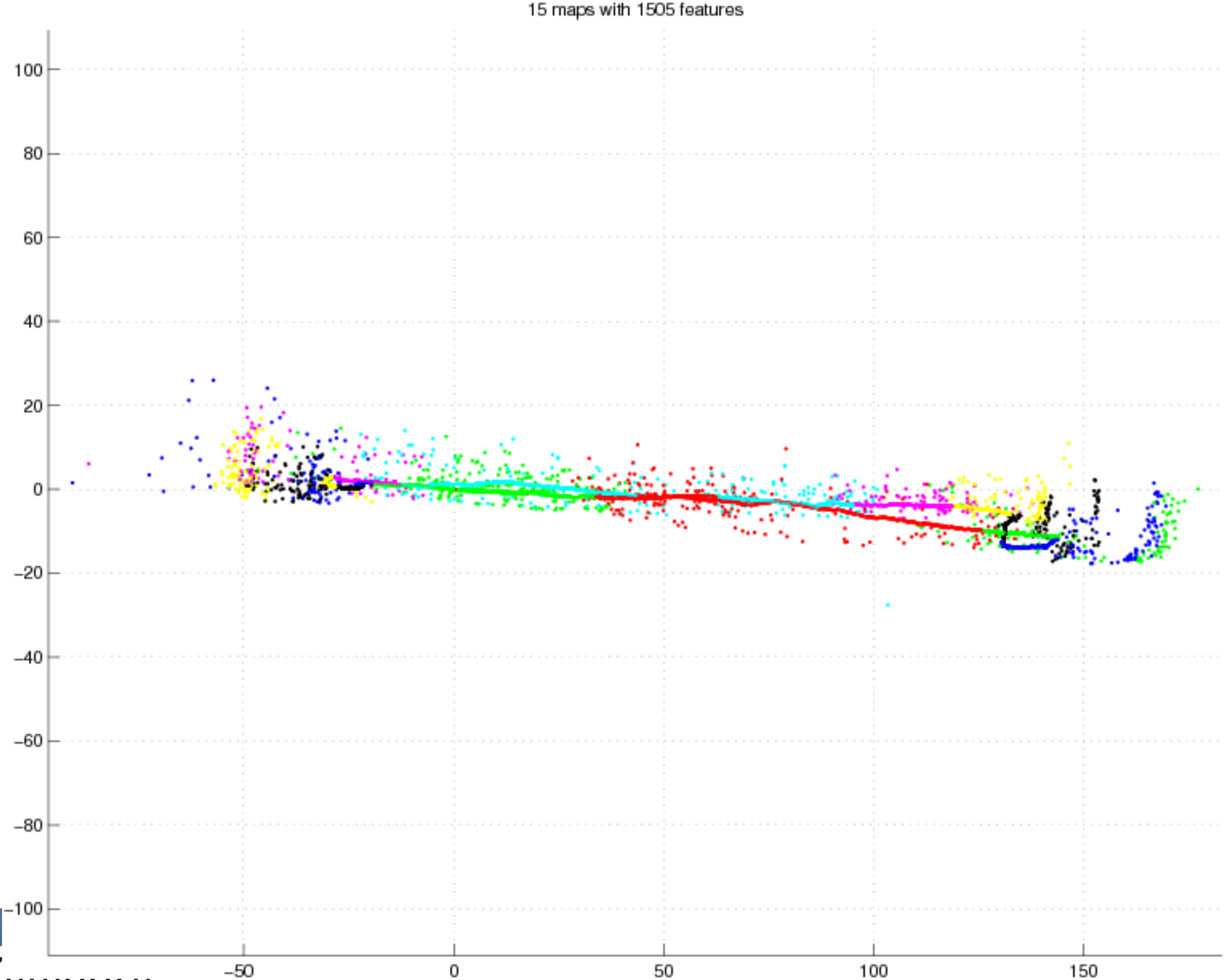
Loop closing: map-to-map matching



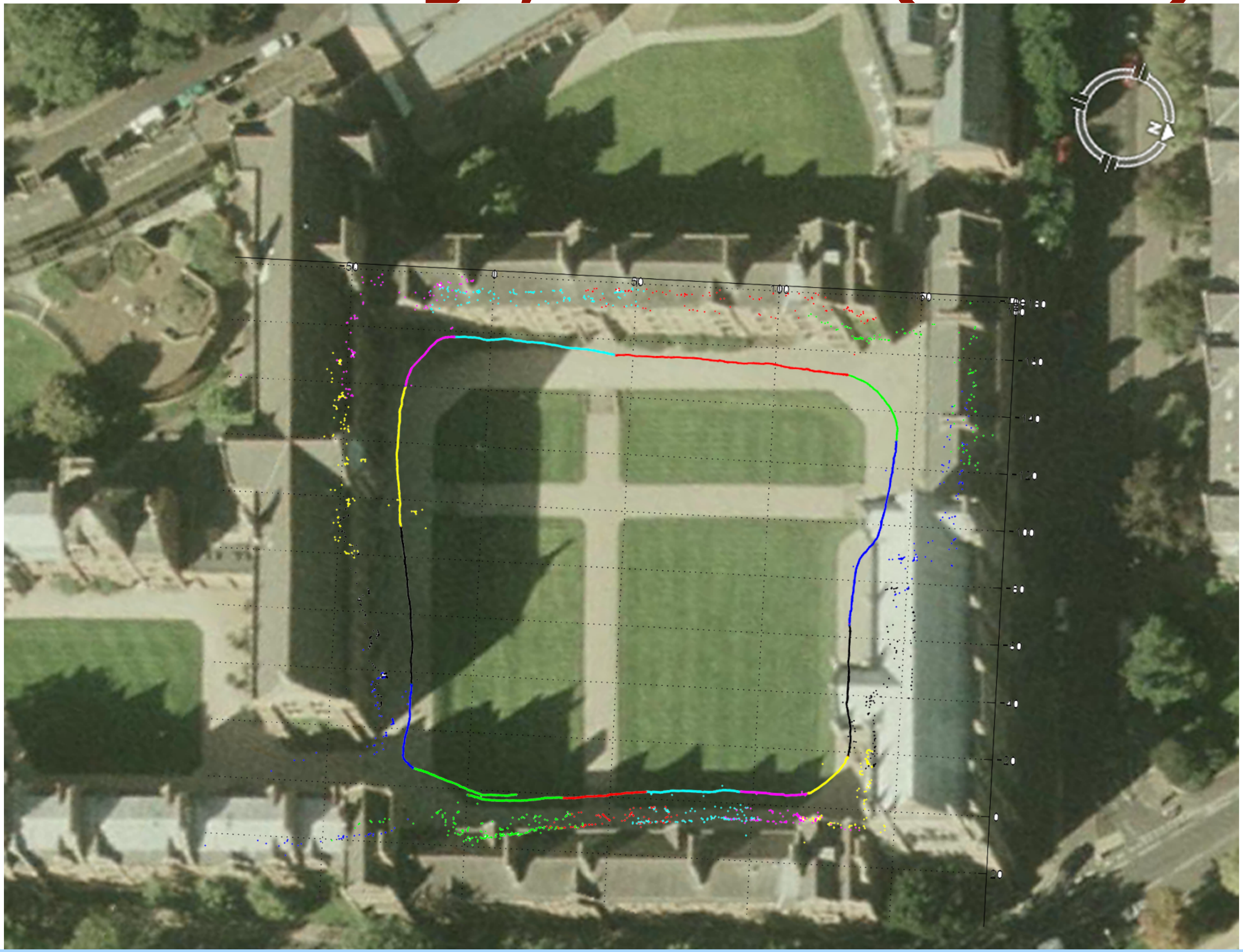
Loop closing



Loop closing (lateral view)



Keble College, Oxford (290m)



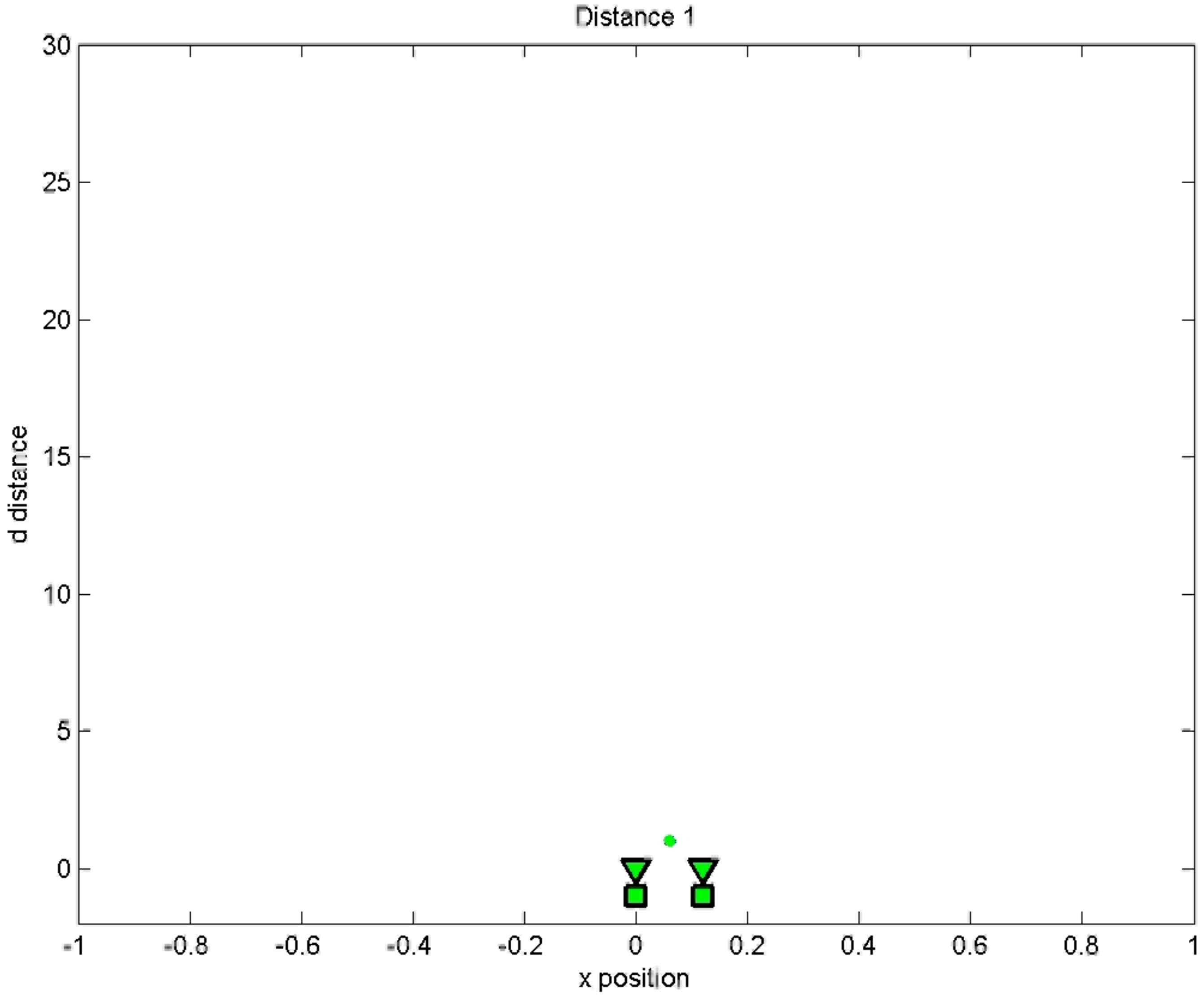
L. Clemente, A. Davison, I. Reid, J. Neira and J.D. Tardós **Mapping Large Loops with a Single Hand-Held Camera**. Robotics: Science and Systems, 2007.

SLAM using only stereo

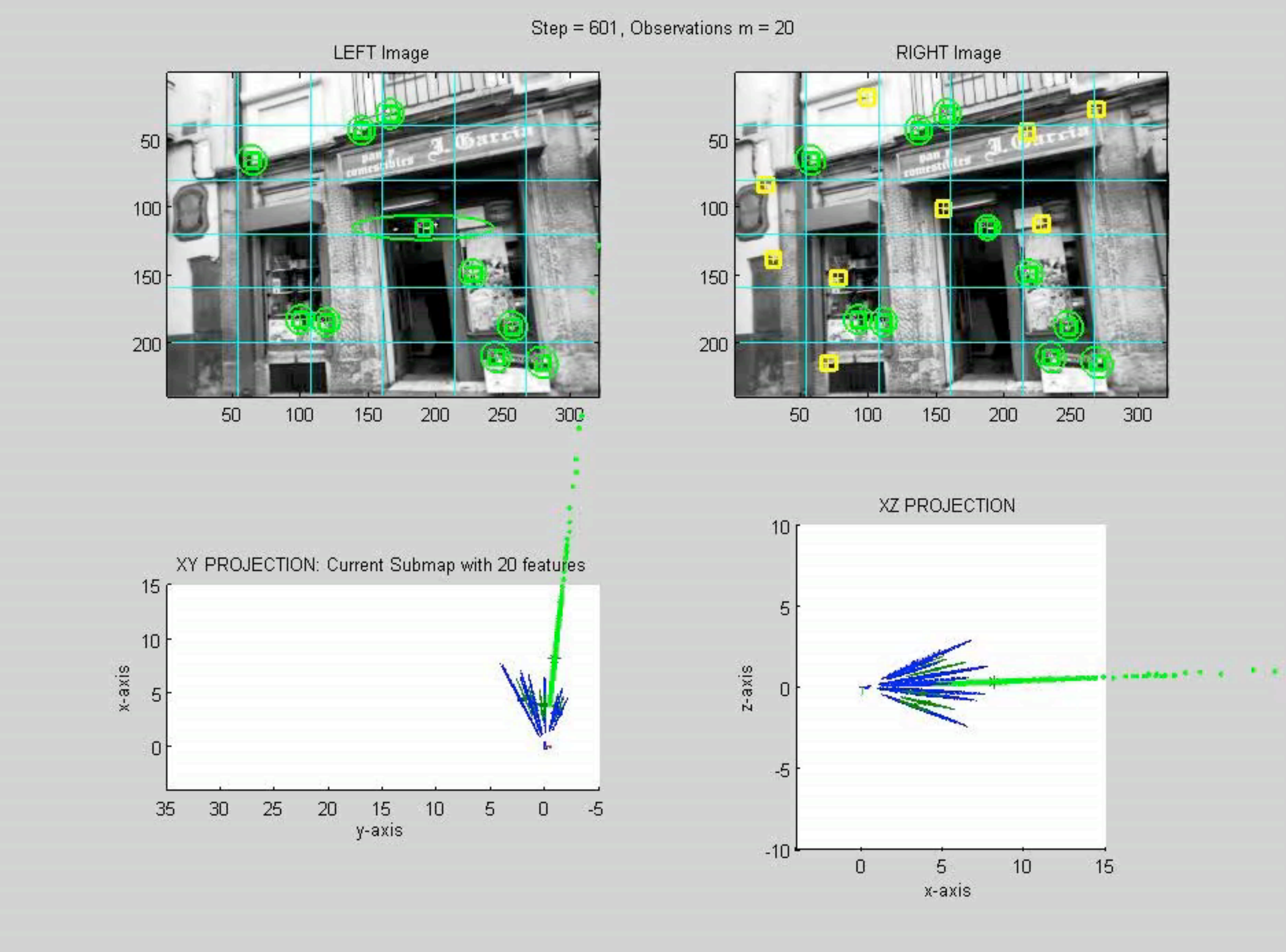
- Experimental setup

A bumblebee, a laptop and a firewire cable

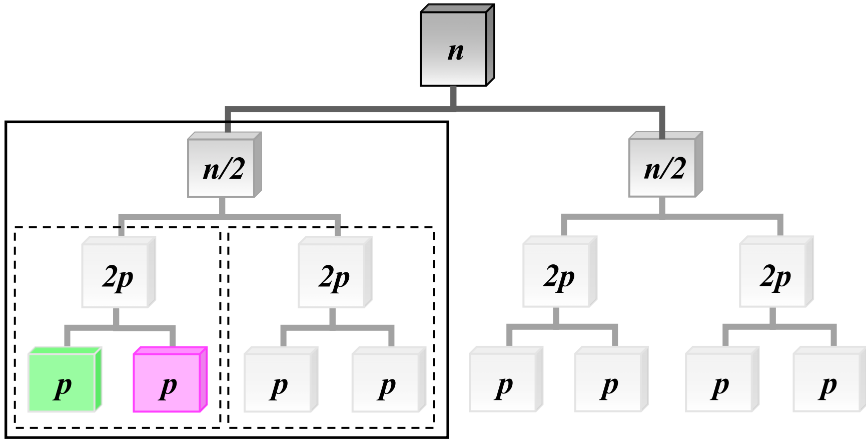
Depth .vs. Inverse Depth

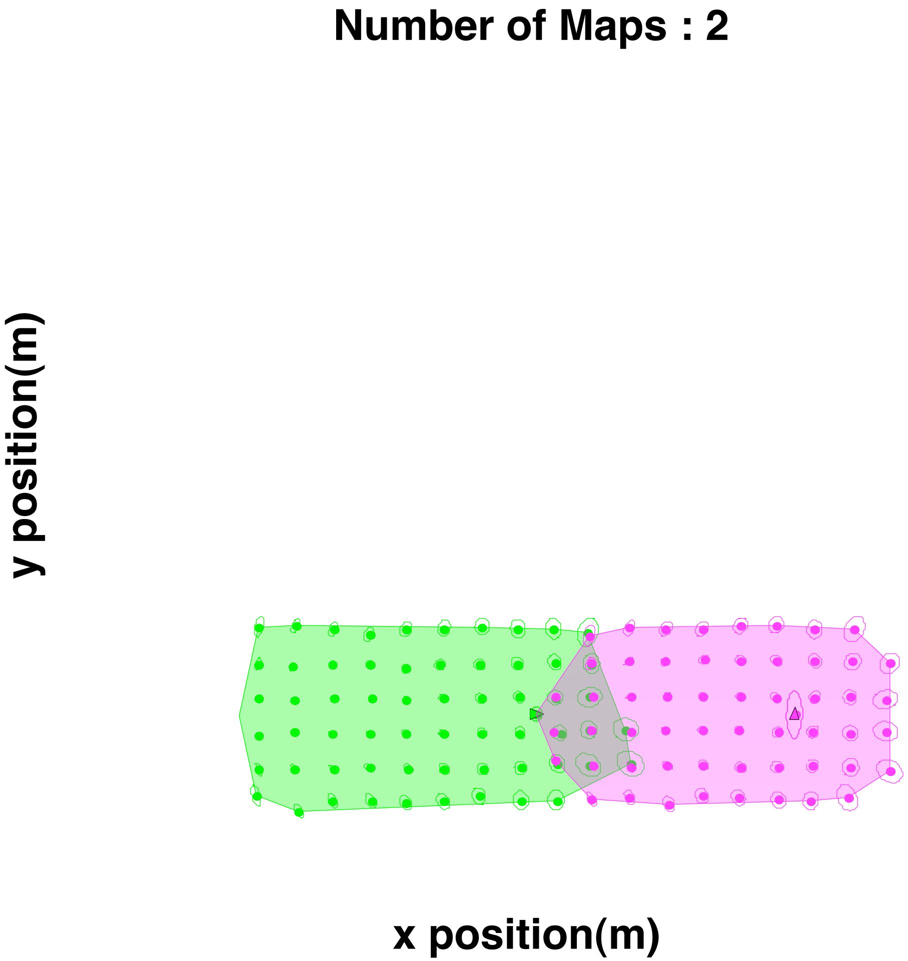


Basic EKF SLAM

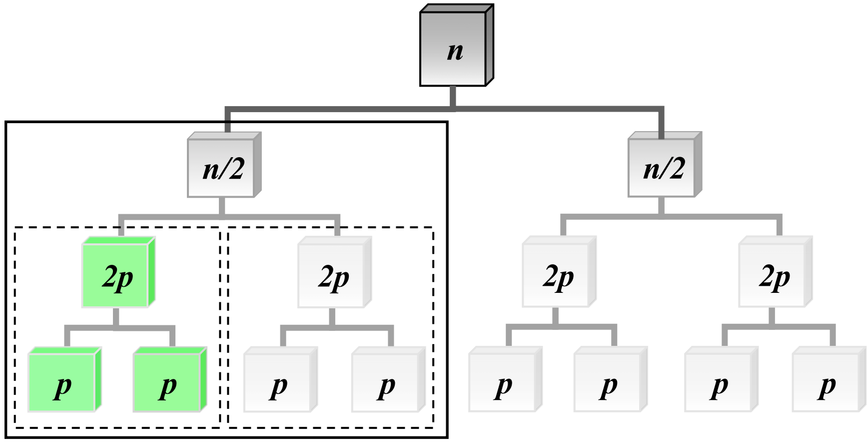


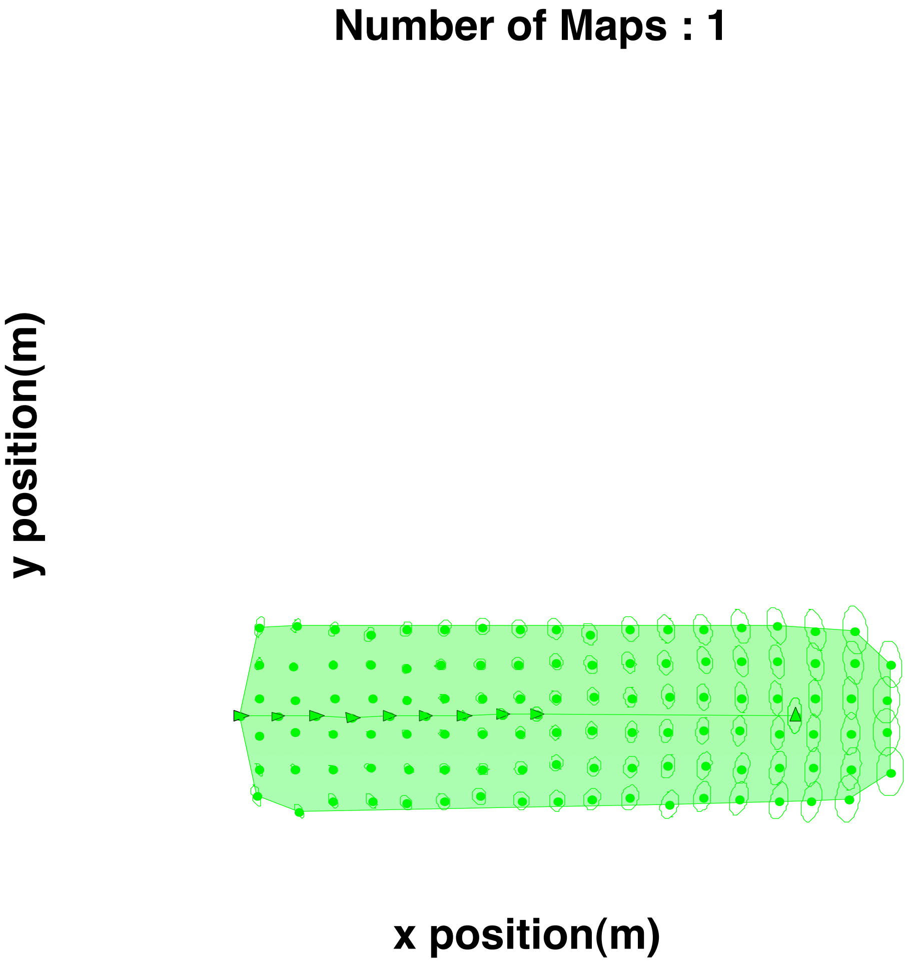
Divide & Conquer SLAM



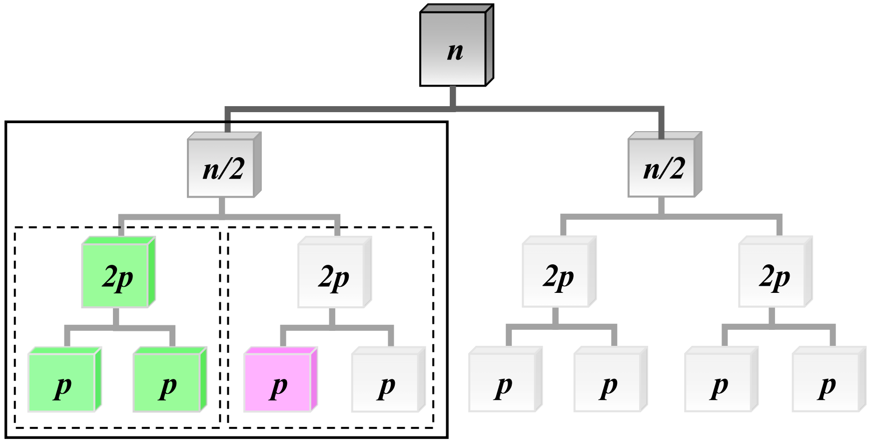


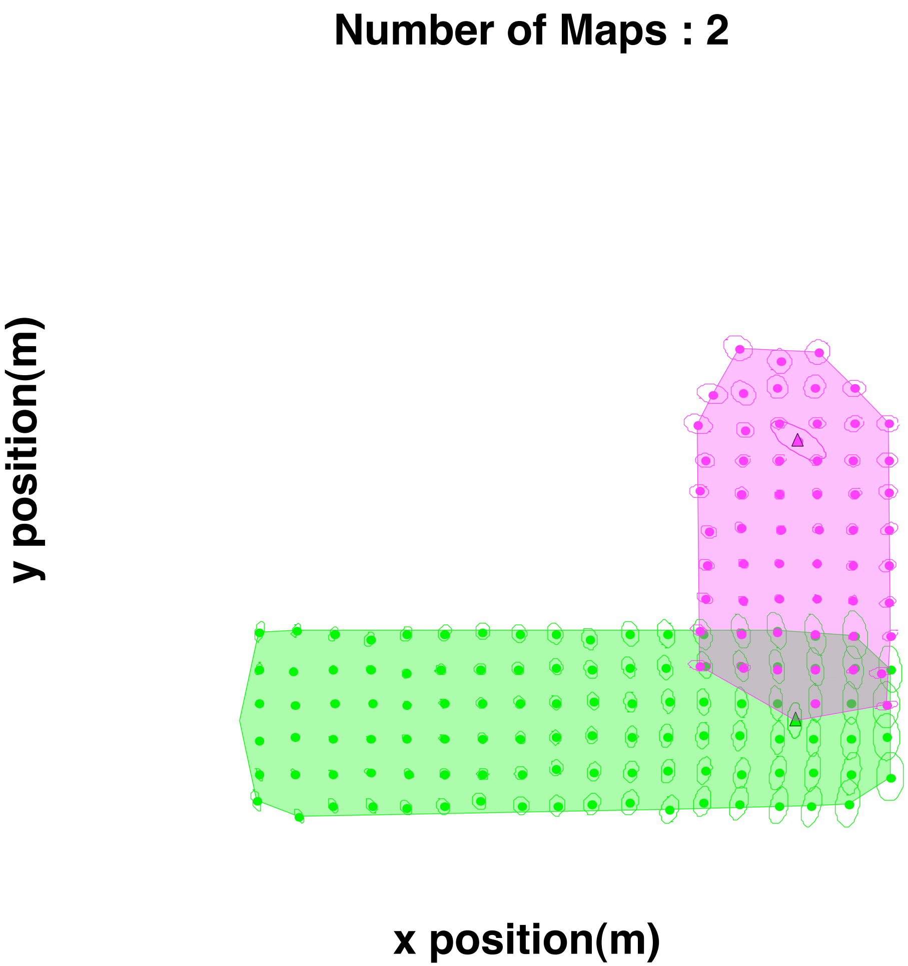
Divide & Conquer SLAM



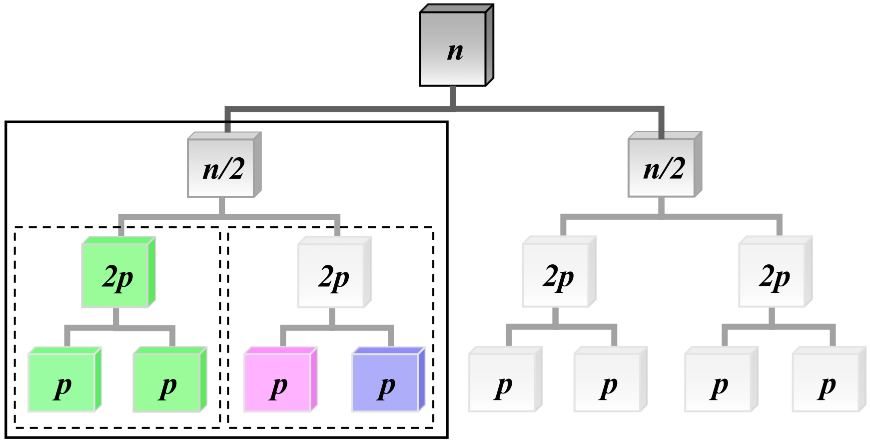


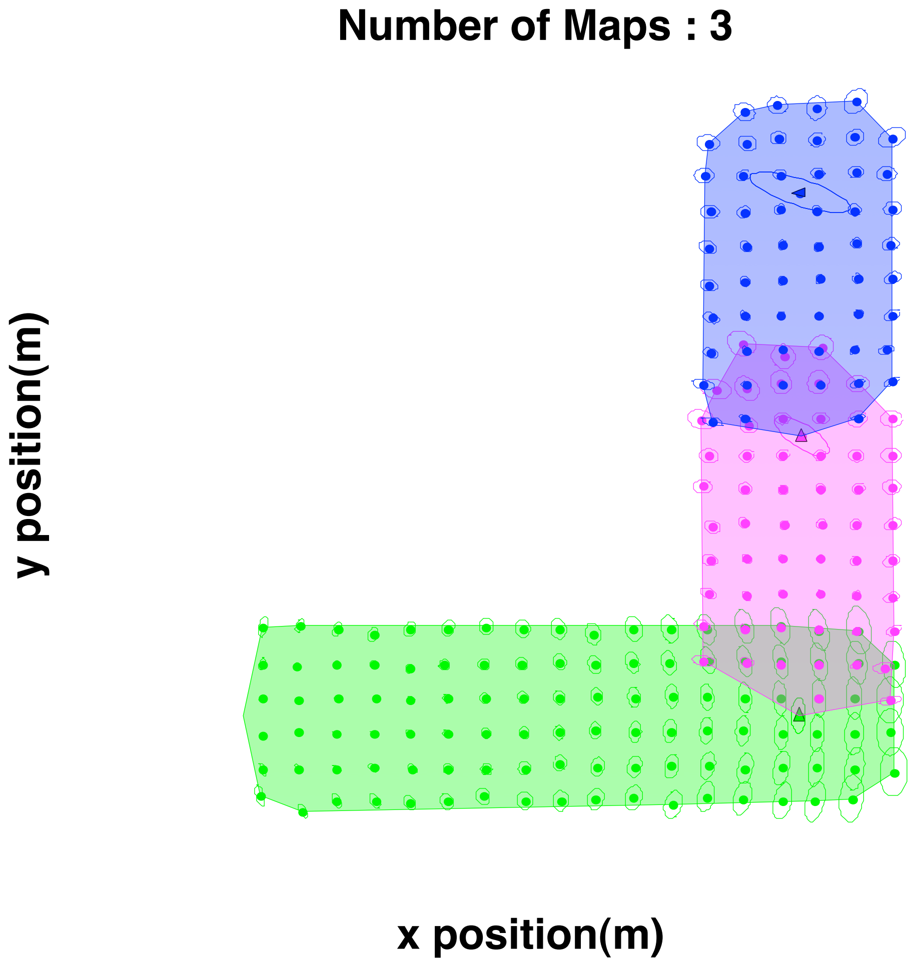
Divide & Conquer SLAM



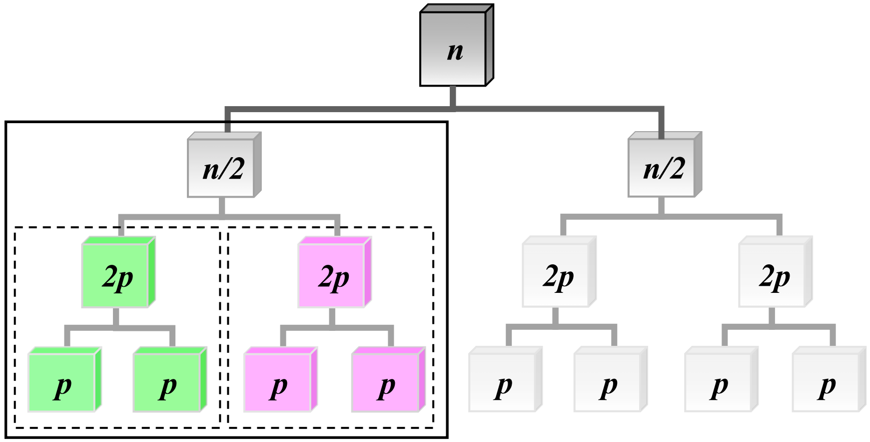


Divide & Conquer SLAM

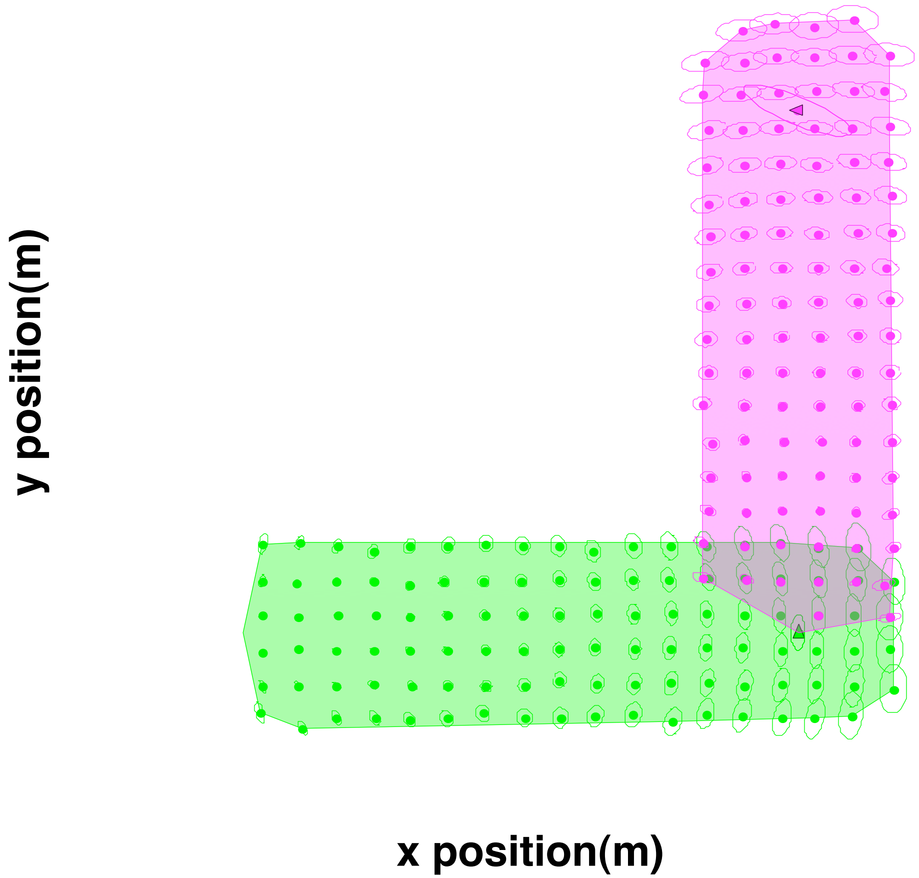




Divide & Conquer SLAM

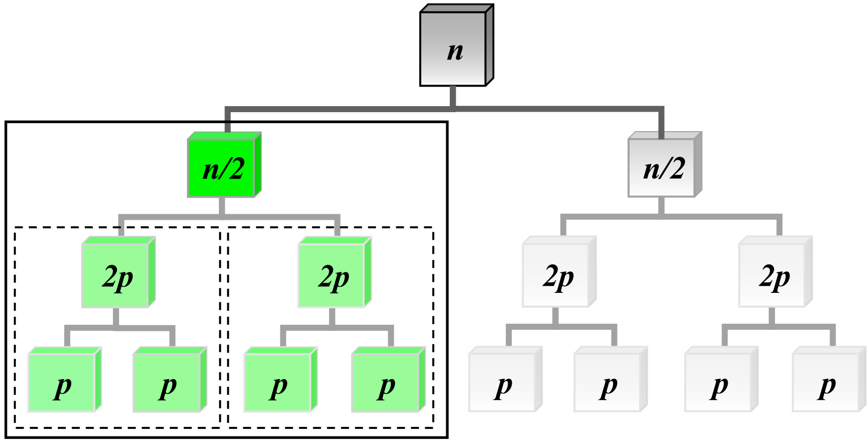


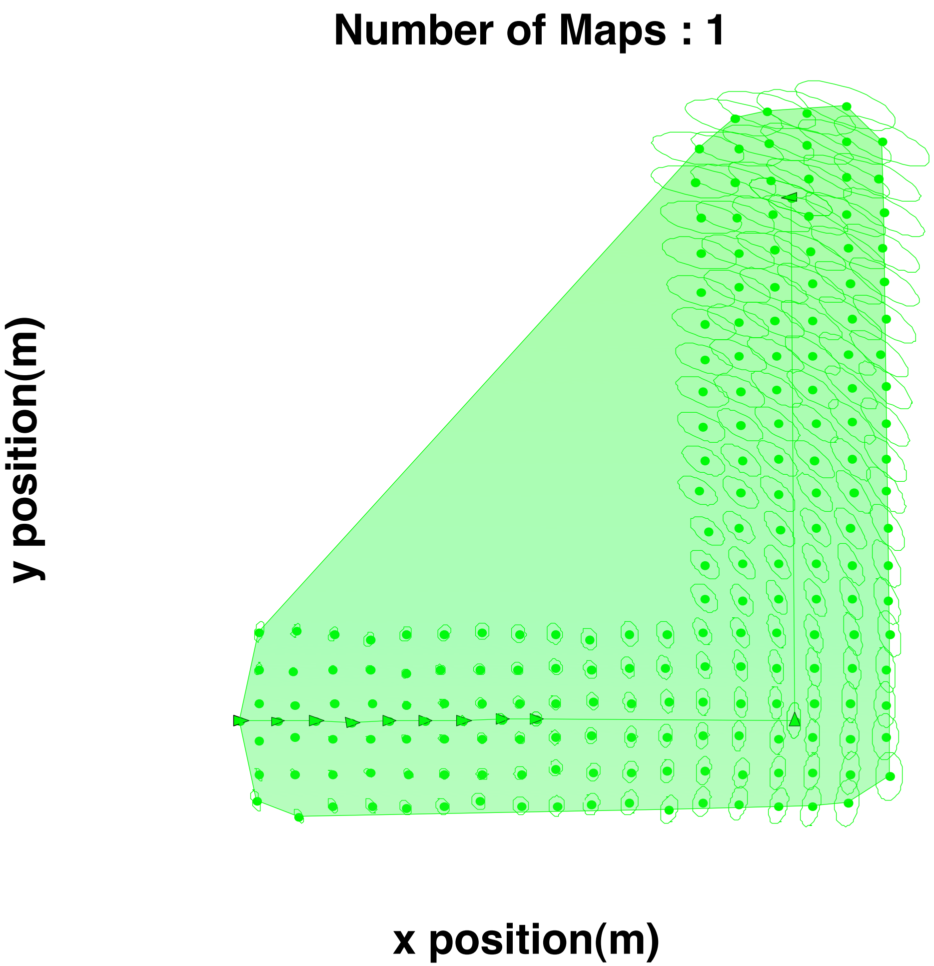
Number of Maps : 2



Universidad Zaragoza 1542

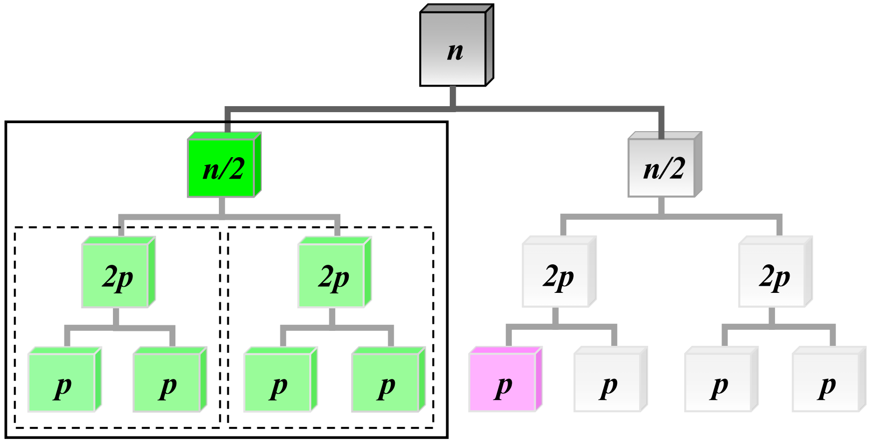
Divide & Conquer SLAM

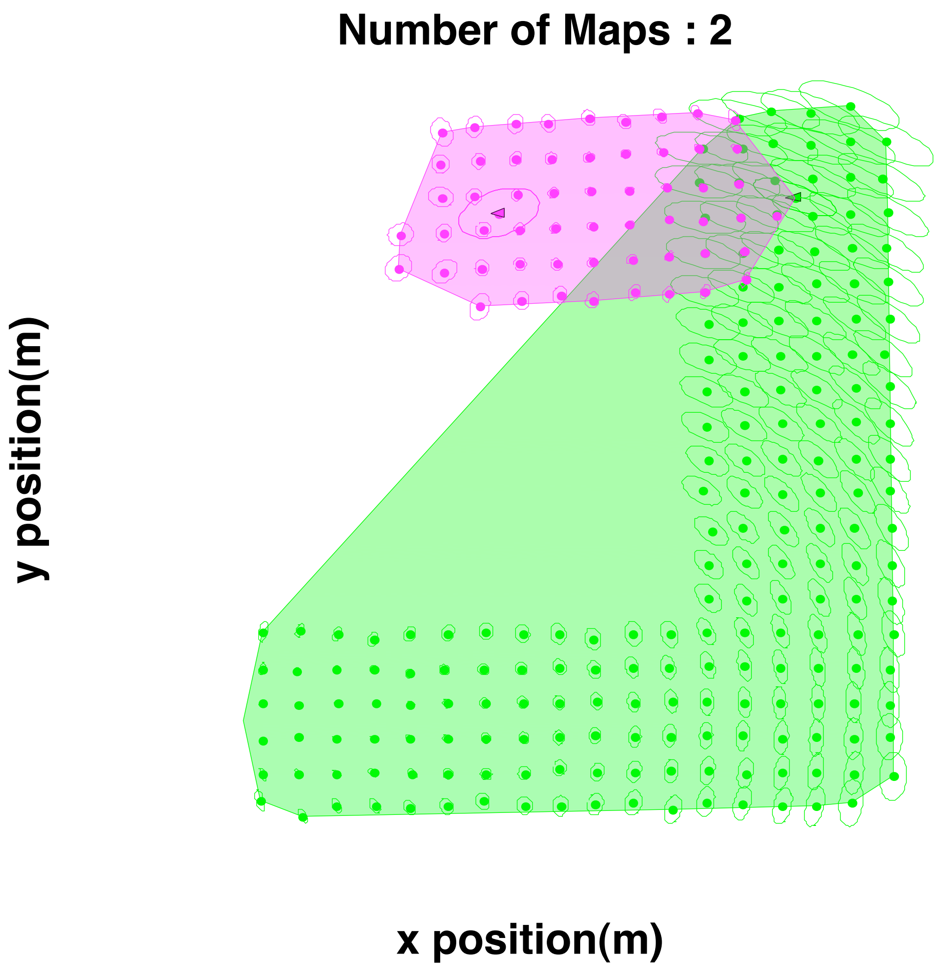




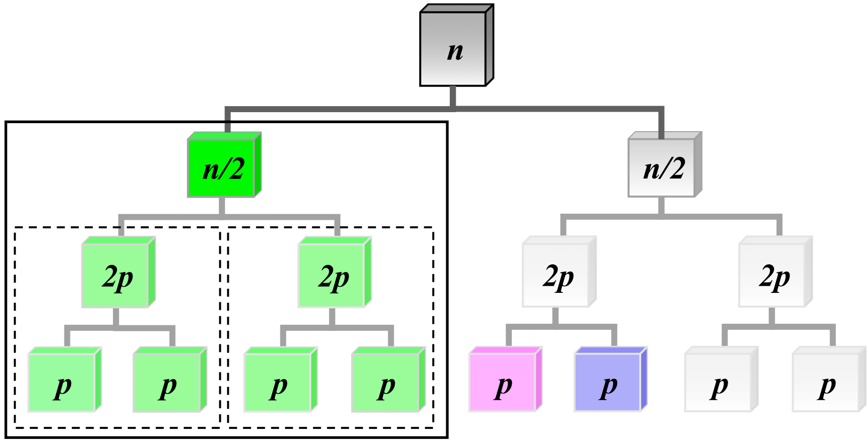
Universidad Zaragoza
1542

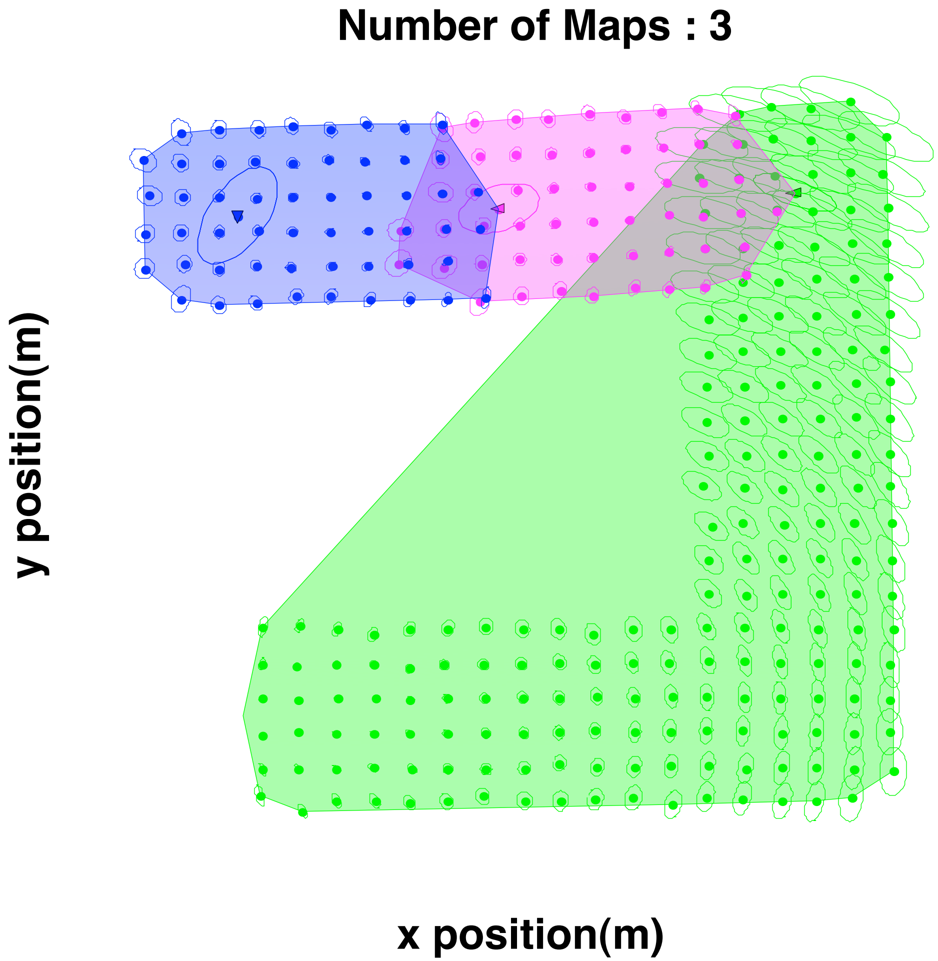
Divide & Conquer SLAM





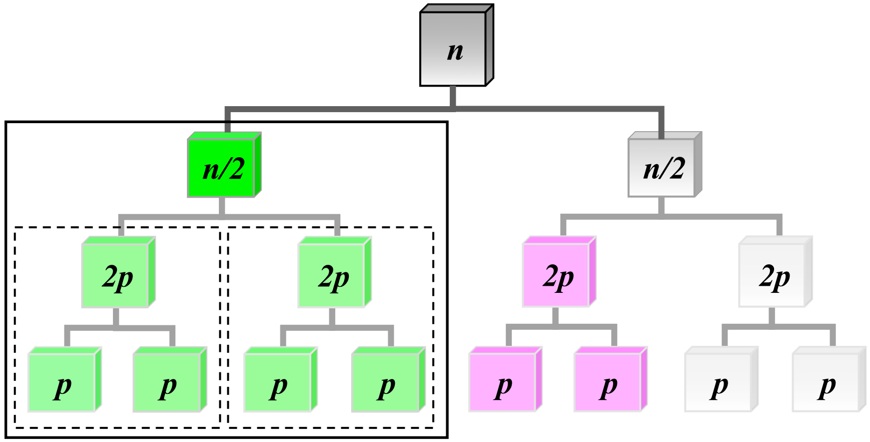
Divide & Conquer SLAM

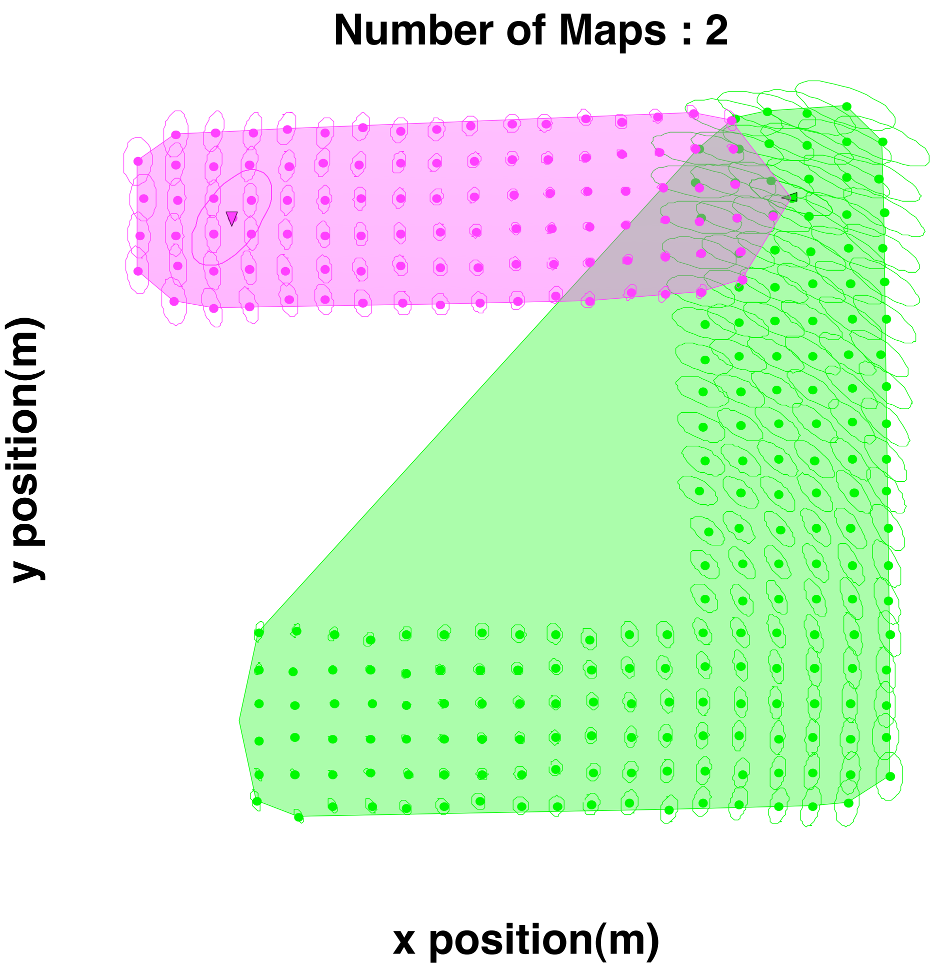




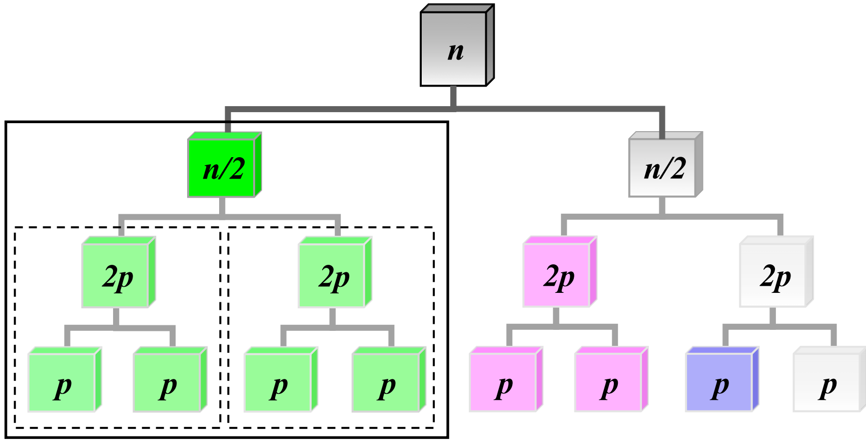
Universidad Zaragoza 1542

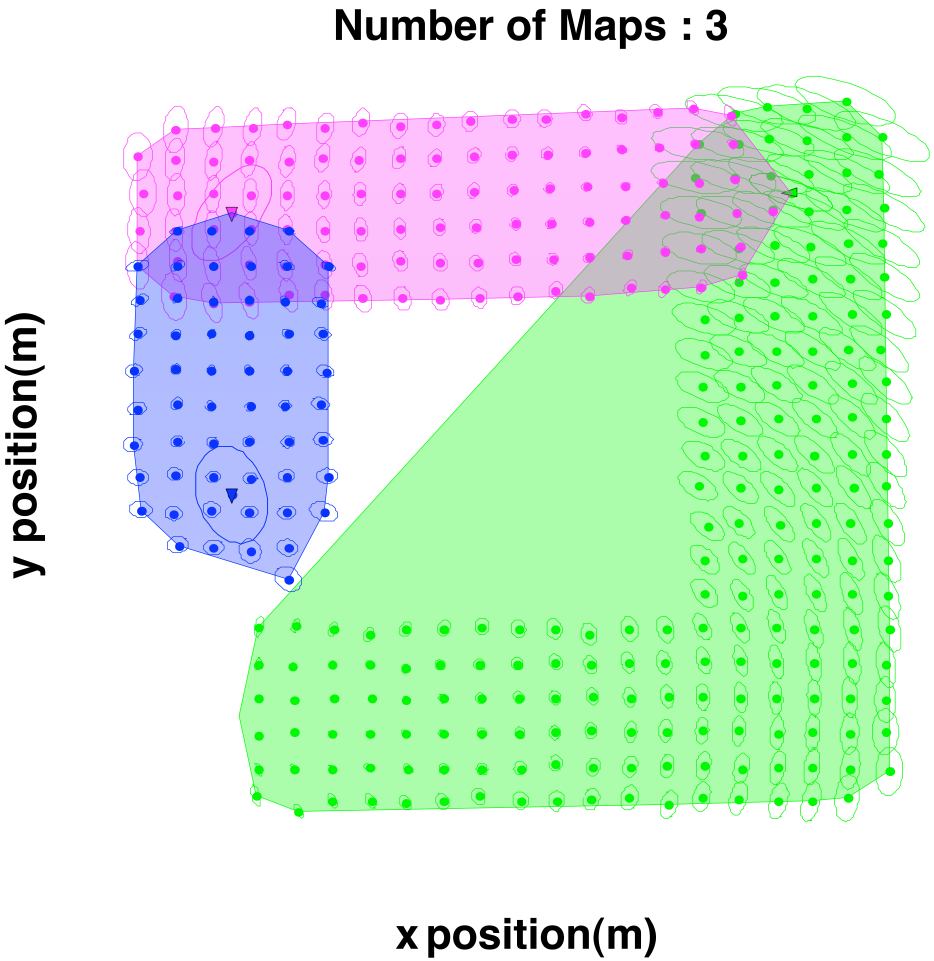
Divide & Conquer SLAM





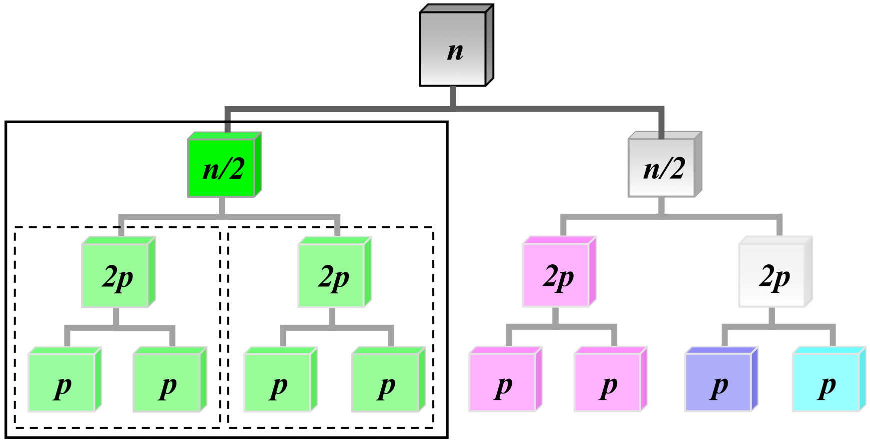
Divide & Conquer SLAM

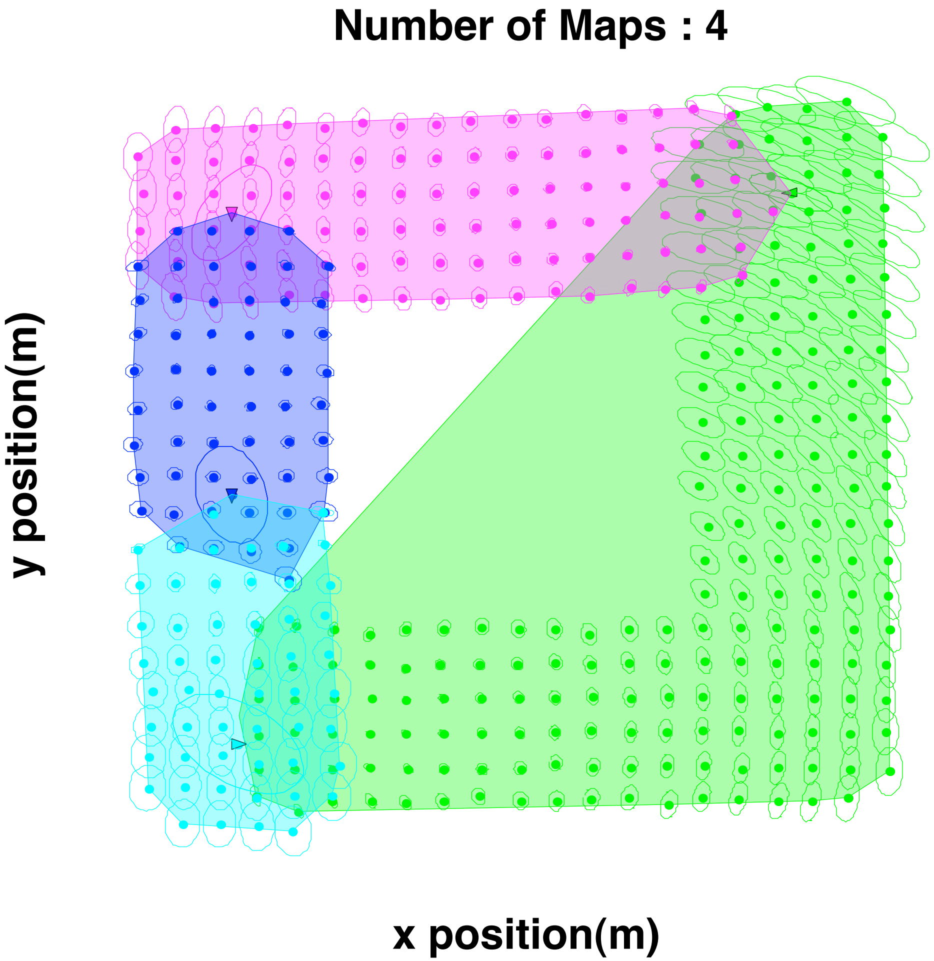




Universidad Zaragoza
1542

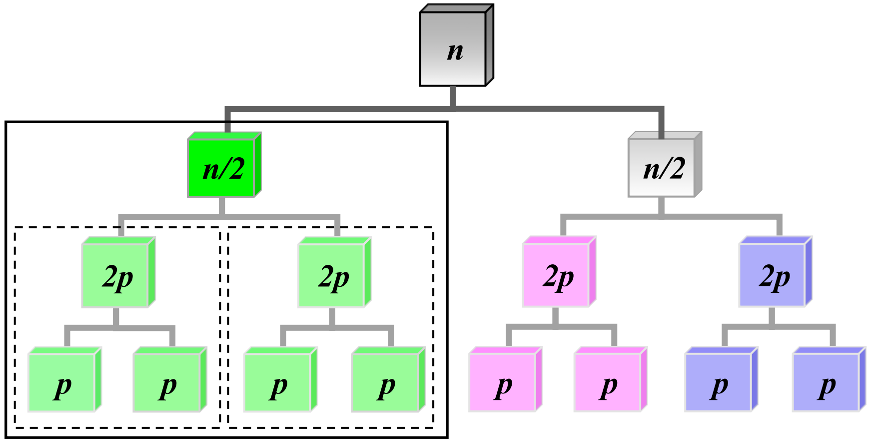
Divide & Conquer SLAM

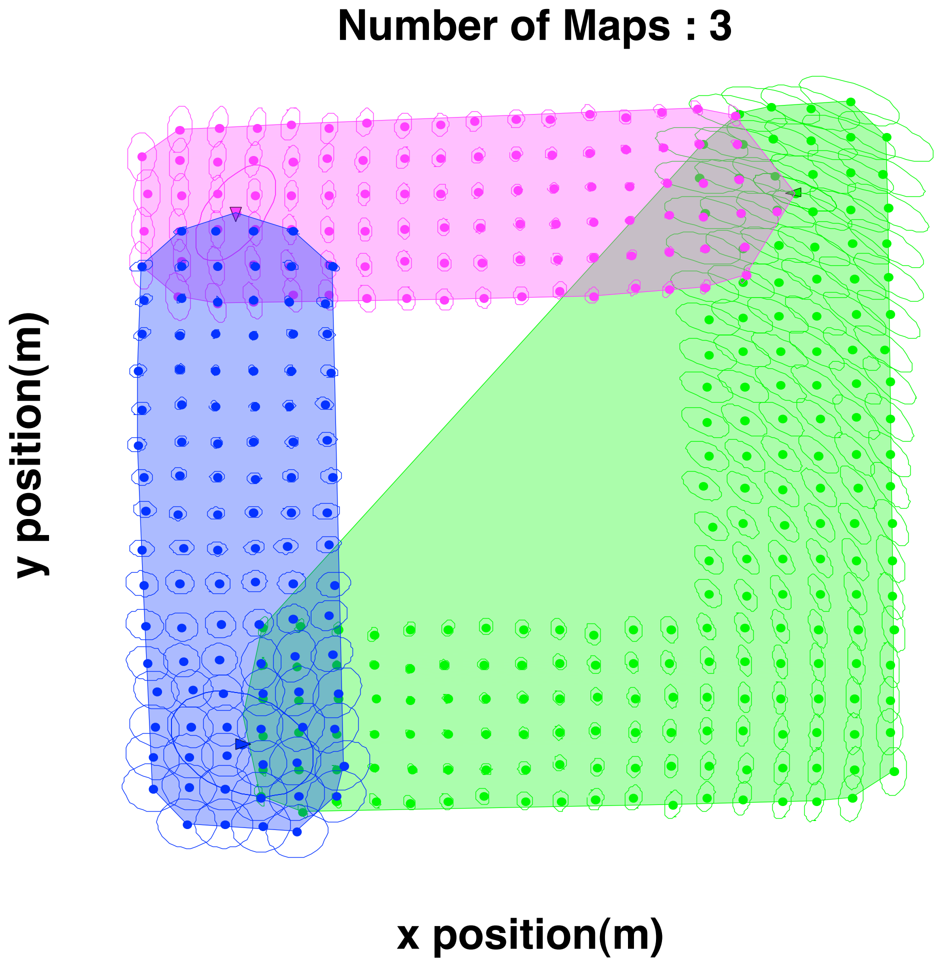




Universidad Zaragoza 1542

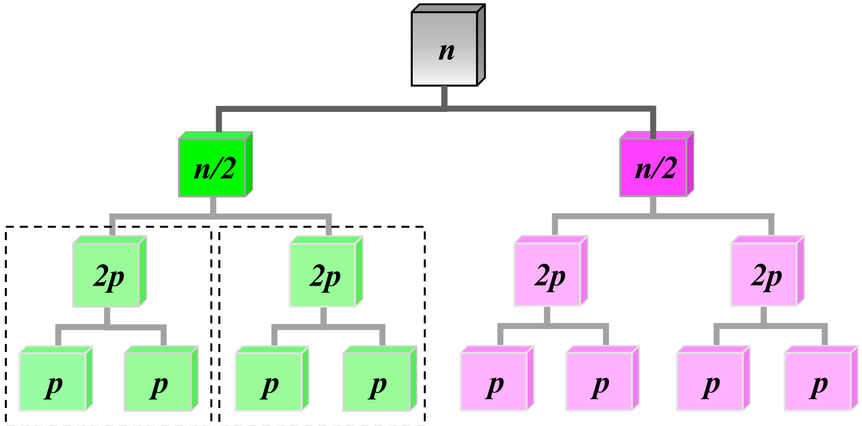
Divide & Conquer SLAM

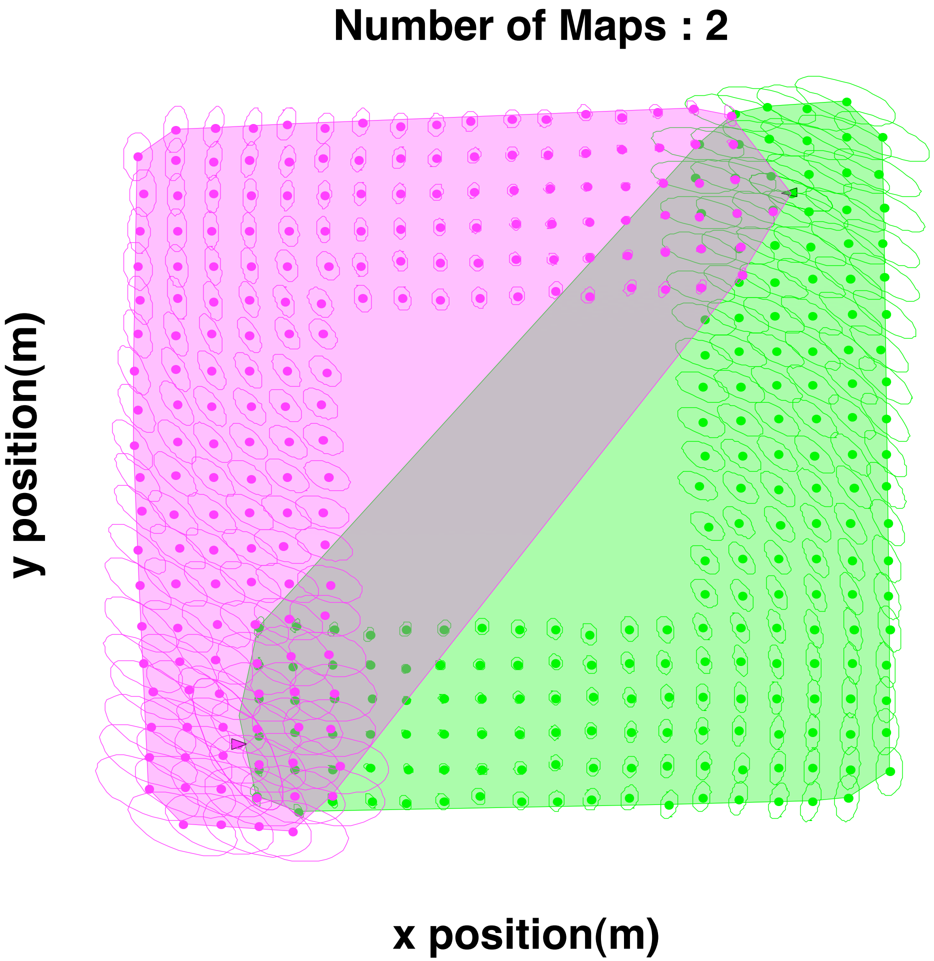




Universidad Zaragoza
1542

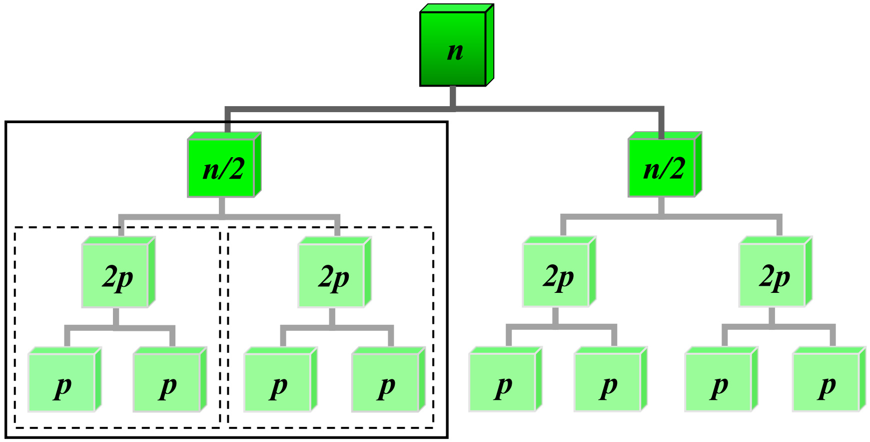
Divide & Conquer SLAM

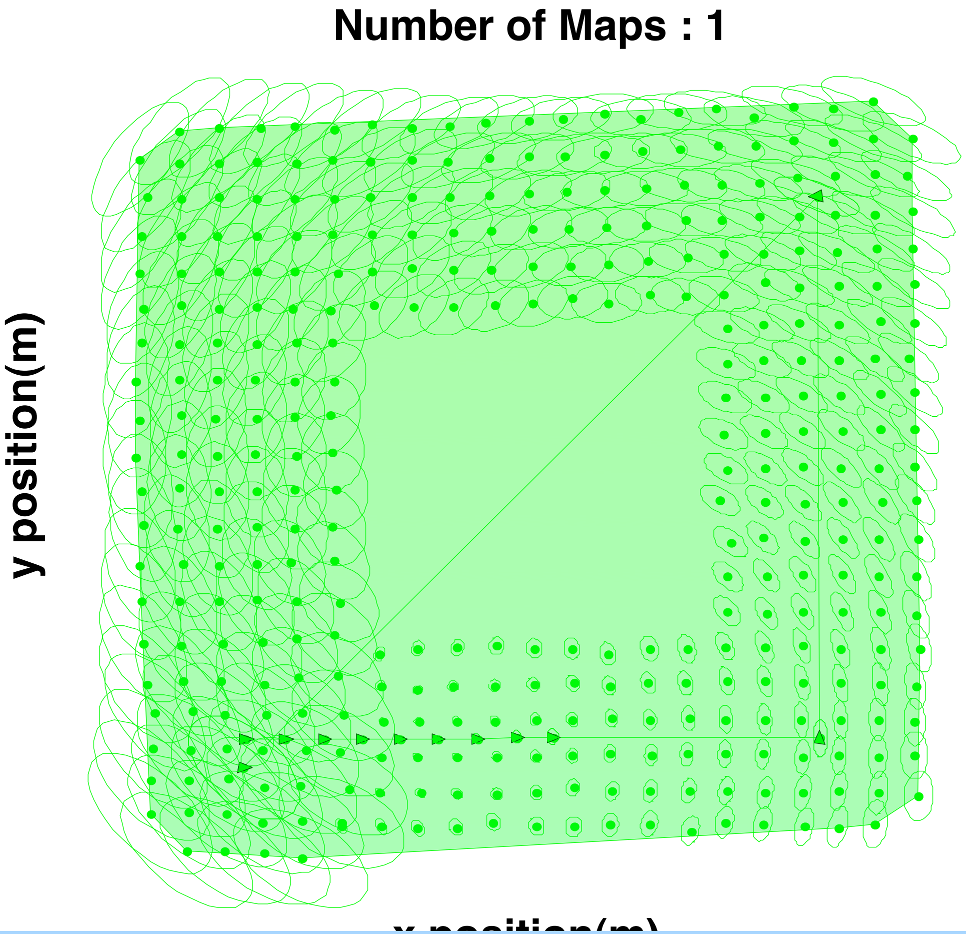




Universidad Zaragoza 1542

Divide & Conquer SLAM

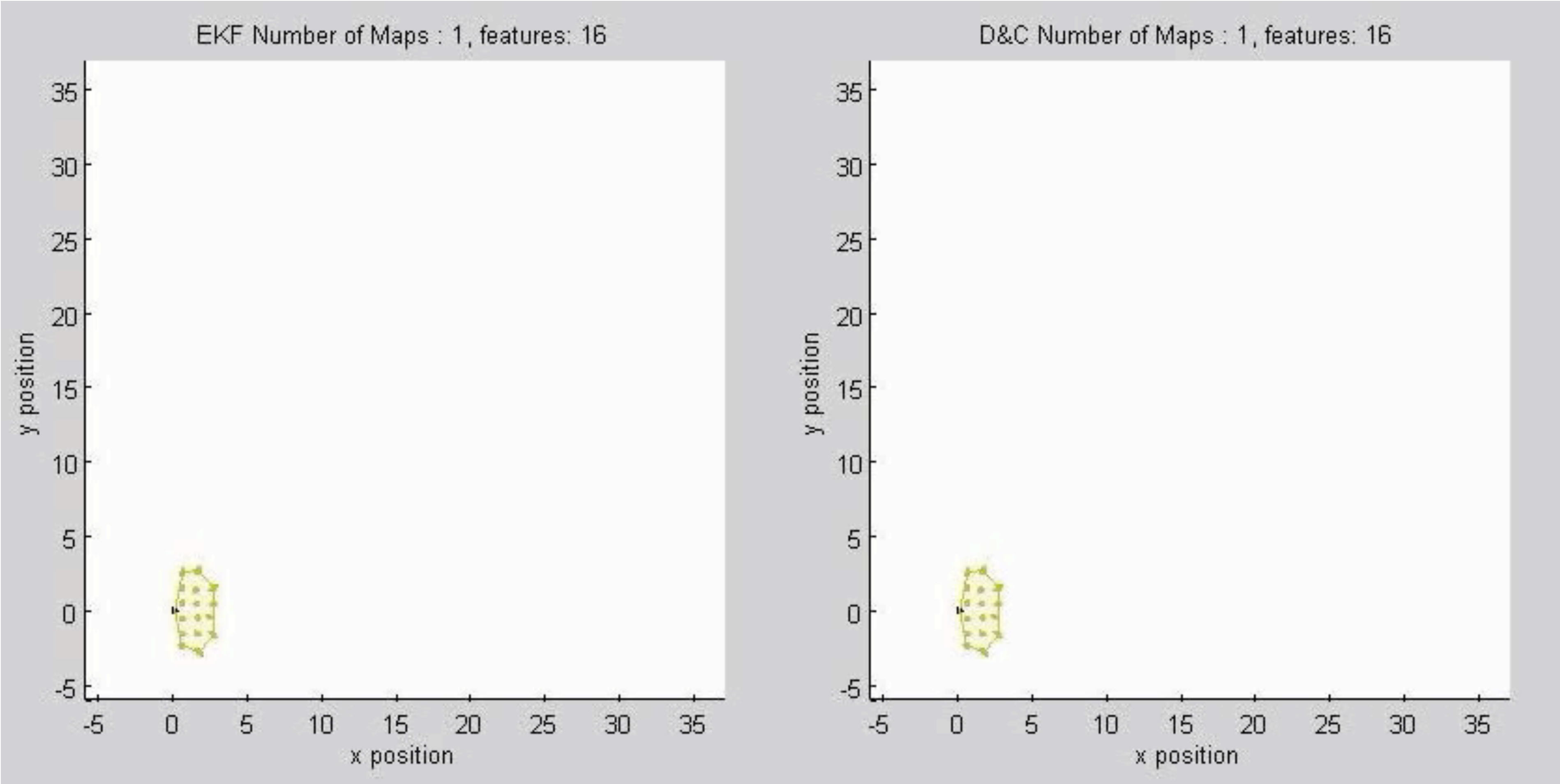




L.M. Paz, P. Jensfelt, J.D. Tardós and J. Neira. **EKF SLAM updates in O(n) with Divide and Conquer SLAM** 2007 IEEE Int. Conf. Robotics and Automation, April 10-14, Rome, Italy

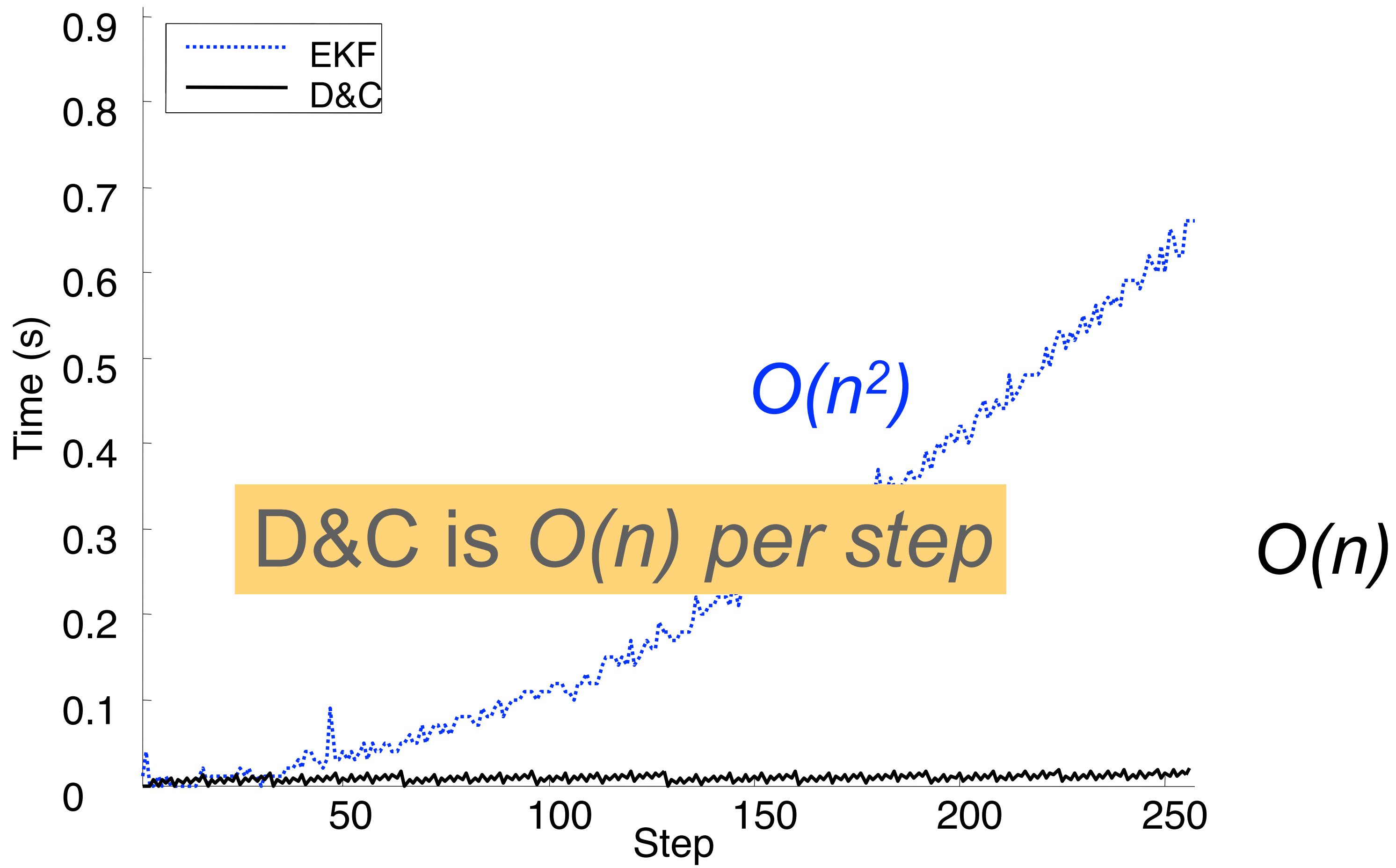
Universidad Zaragoza

Loop Trajectory

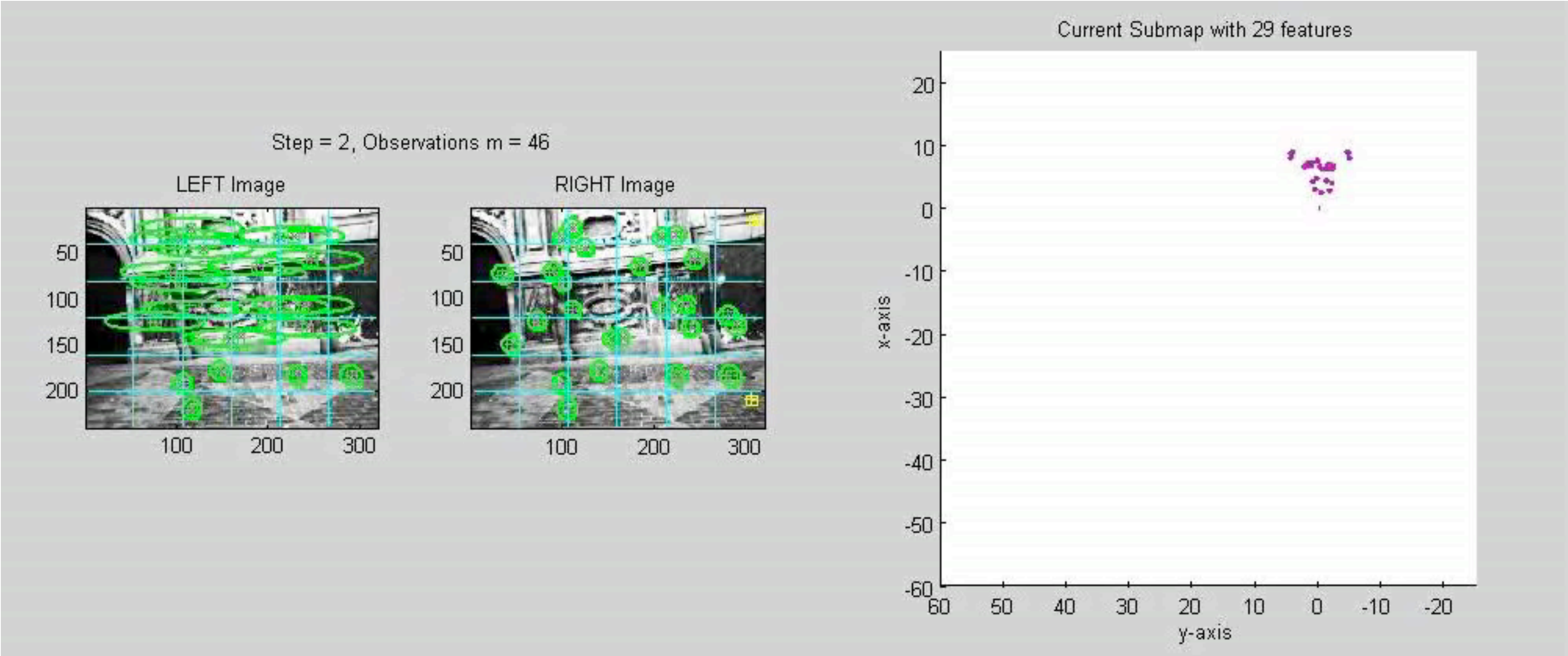


L. Paz, J. Neira and J.D. Tardós **Divide and Conquer: EKF SLAM in O(n)**. IEEE Transactions on Robotics, October 2008.

Amortized cost per step



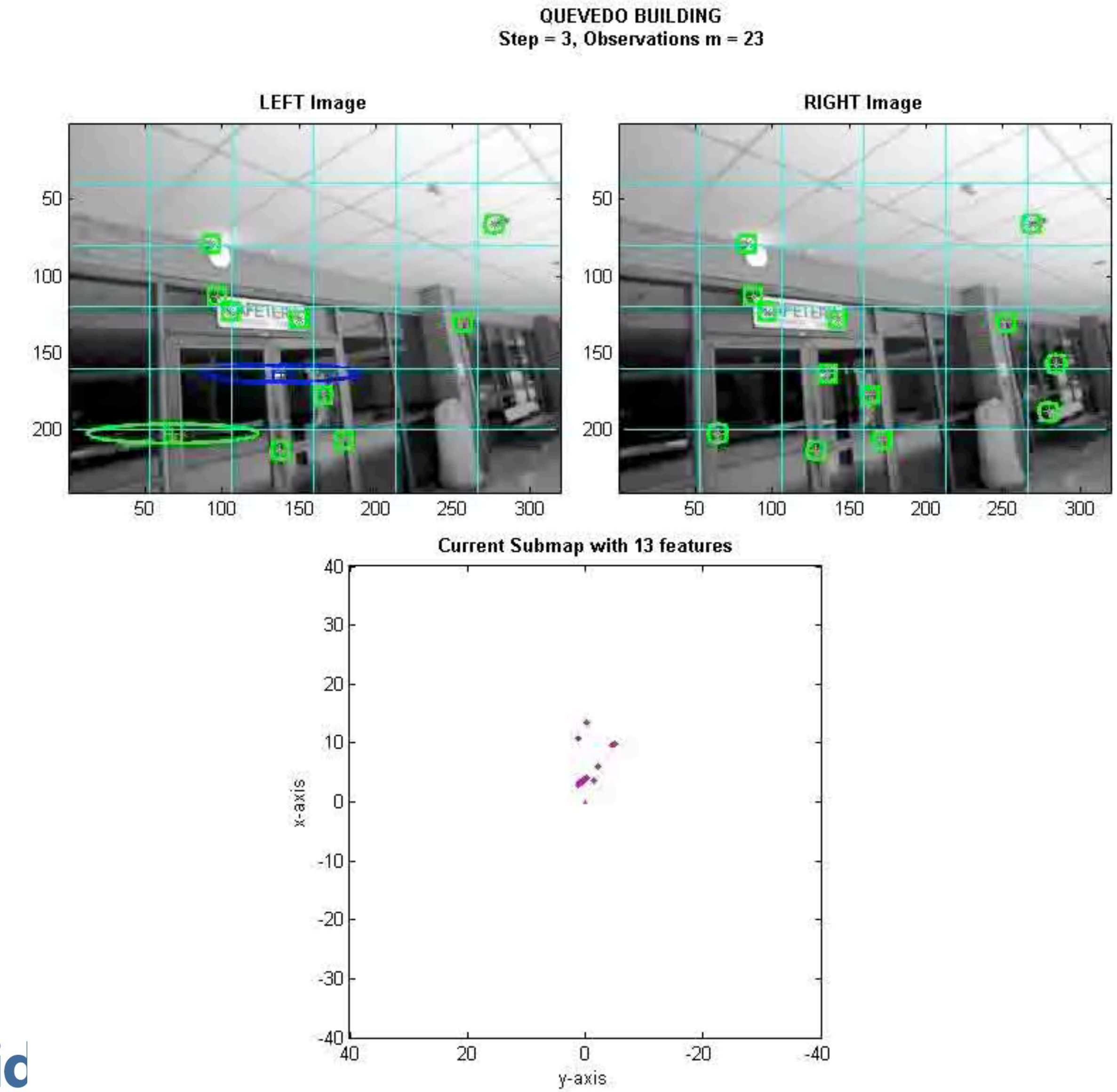
6DOF SLAM with stereo



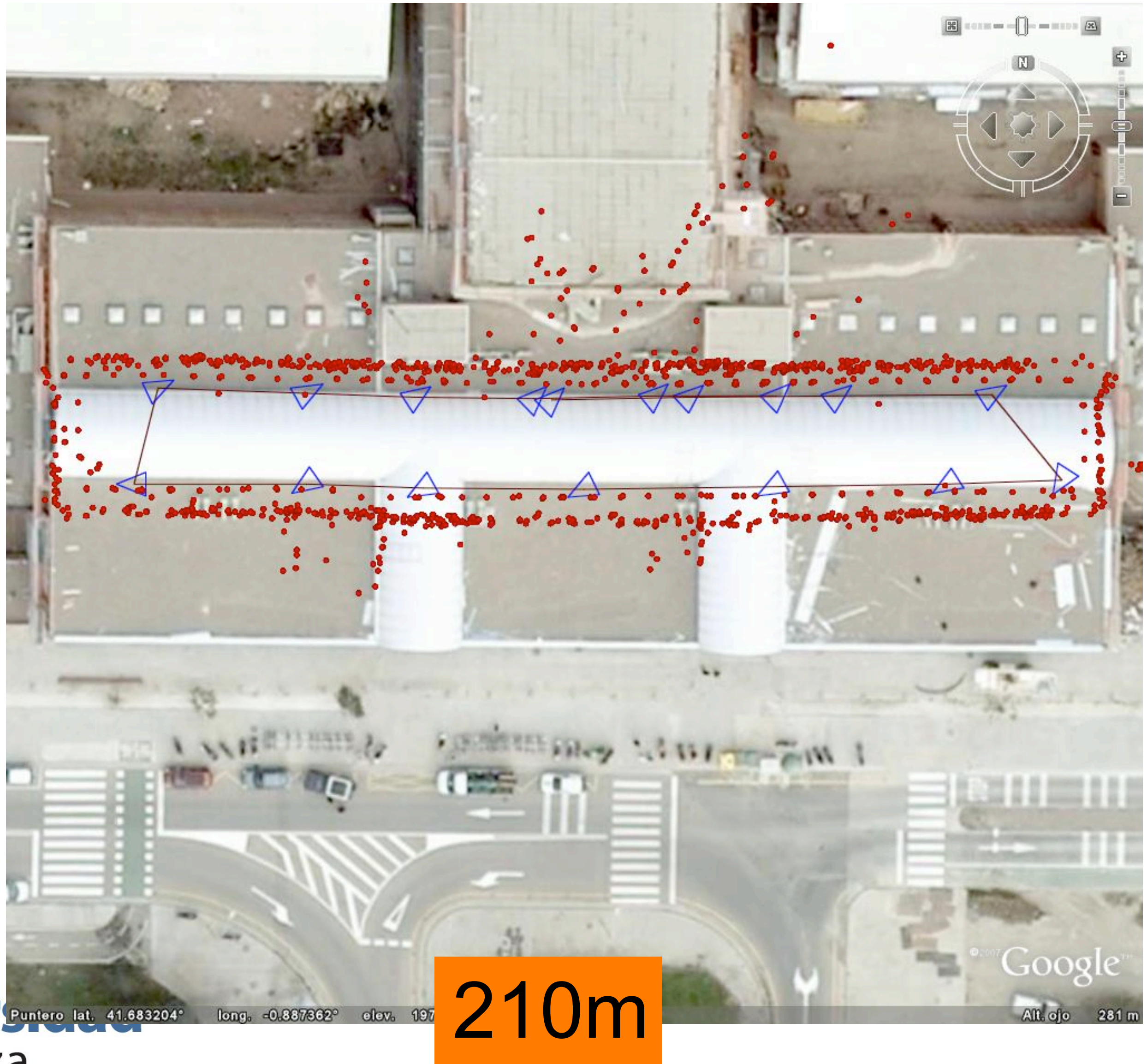
L. Paz, P. Piníes, J. Neira and J.D. Tardós **Large Scale 6DOF SLAM with Stereo-in-Hand.** IEEE Transactions on Robotics, 2008.

6Dof Stereo SLAM, outdoors

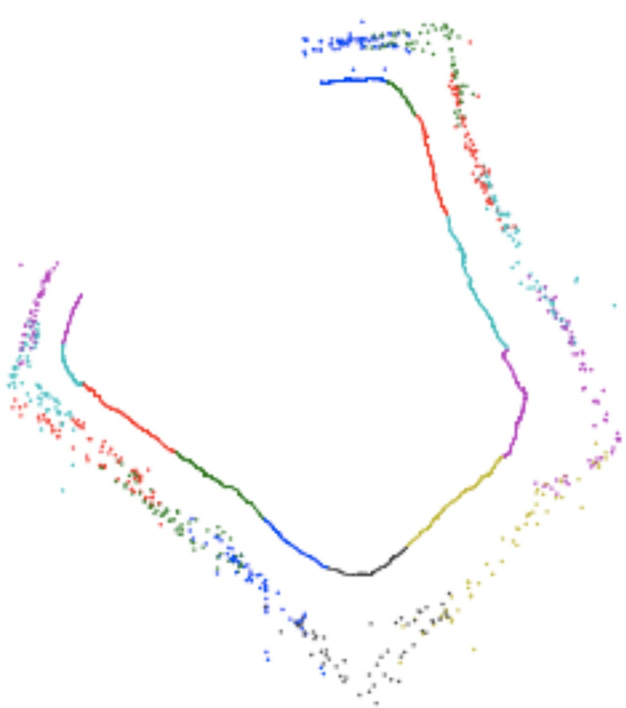
6Dof Stereo SLAM, indoors

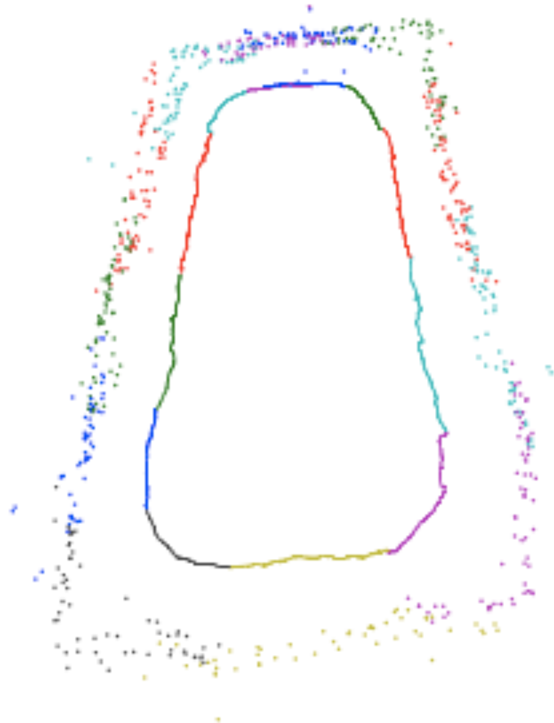


6Dof Stereo SLAM, indoors



The Loop Closing problem

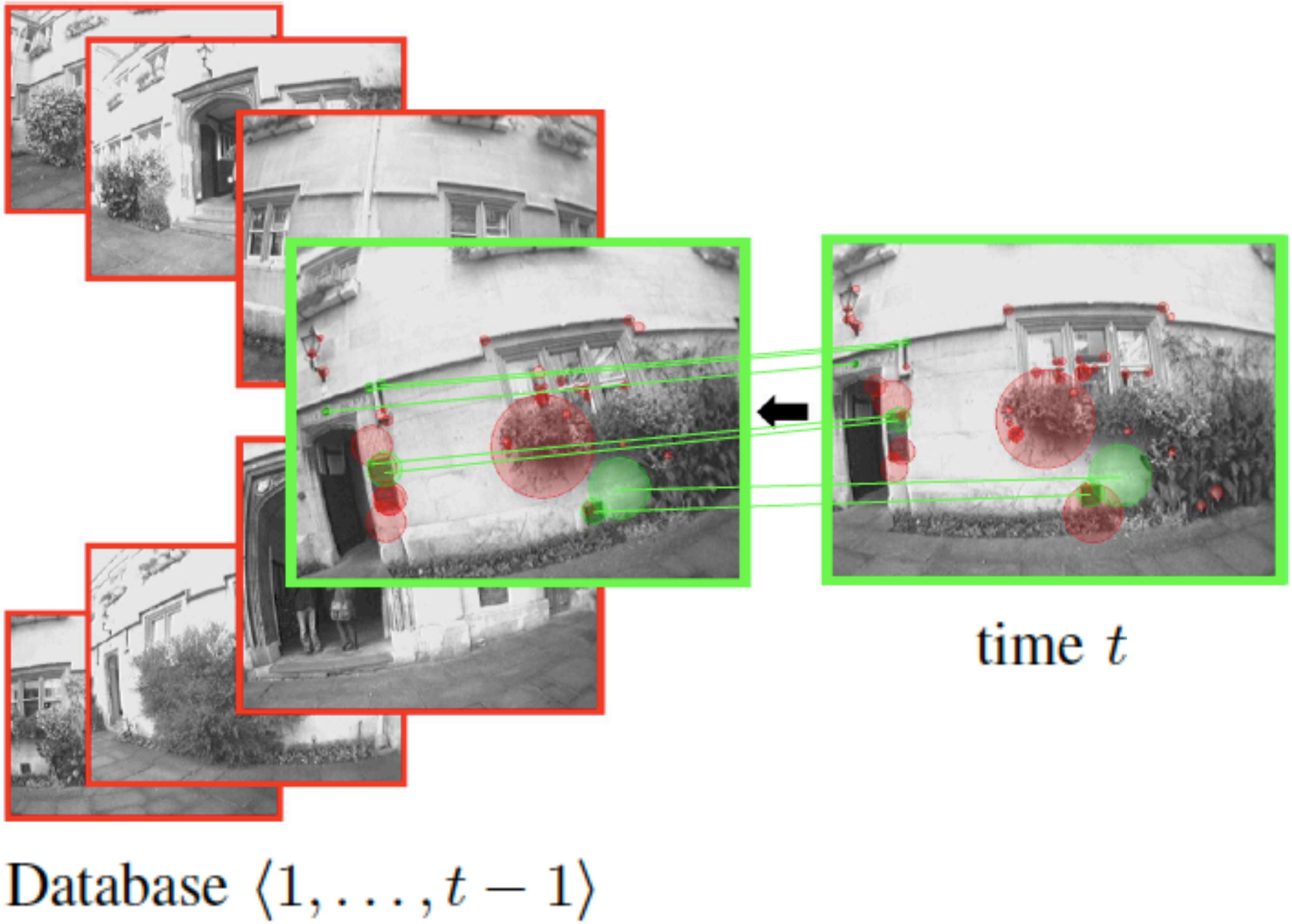




- **Detected loop closings** increment map precision
- **Undetected ones** make the system miss the opportunity to do so
- **Incorrect loop closings** destroy the results

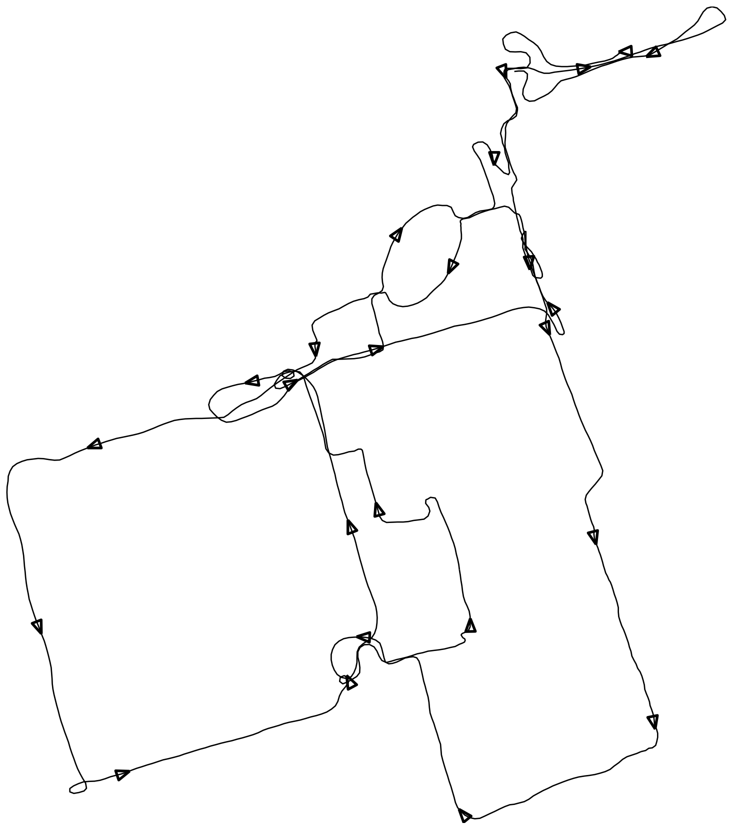
Visual Loop Closing

Input: Scene at time t, Database $\langle 1, \ldots, t-1 \rangle$
Output: Time t' of the revisited place, or null



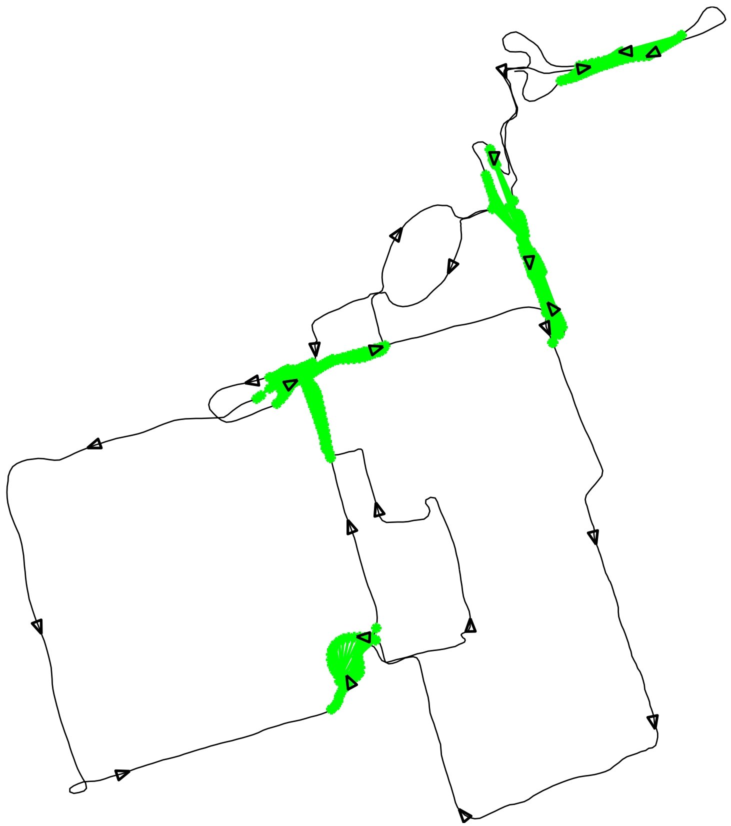
The RAWSEEDS Project

- **Ground Truth trajectory (Outdoors)**



Obtained with GPS

Ground Truth Loop Closings (Outdoors)

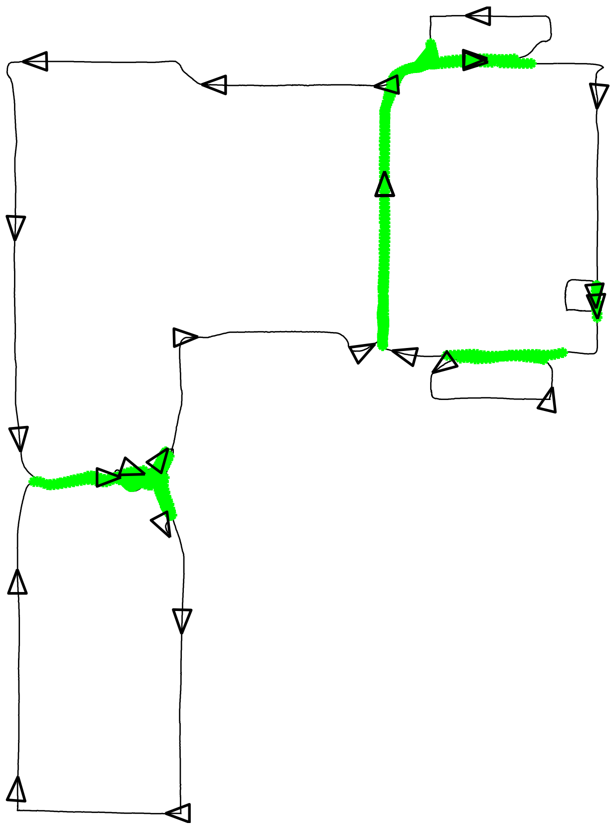


Painstakingly manual GT

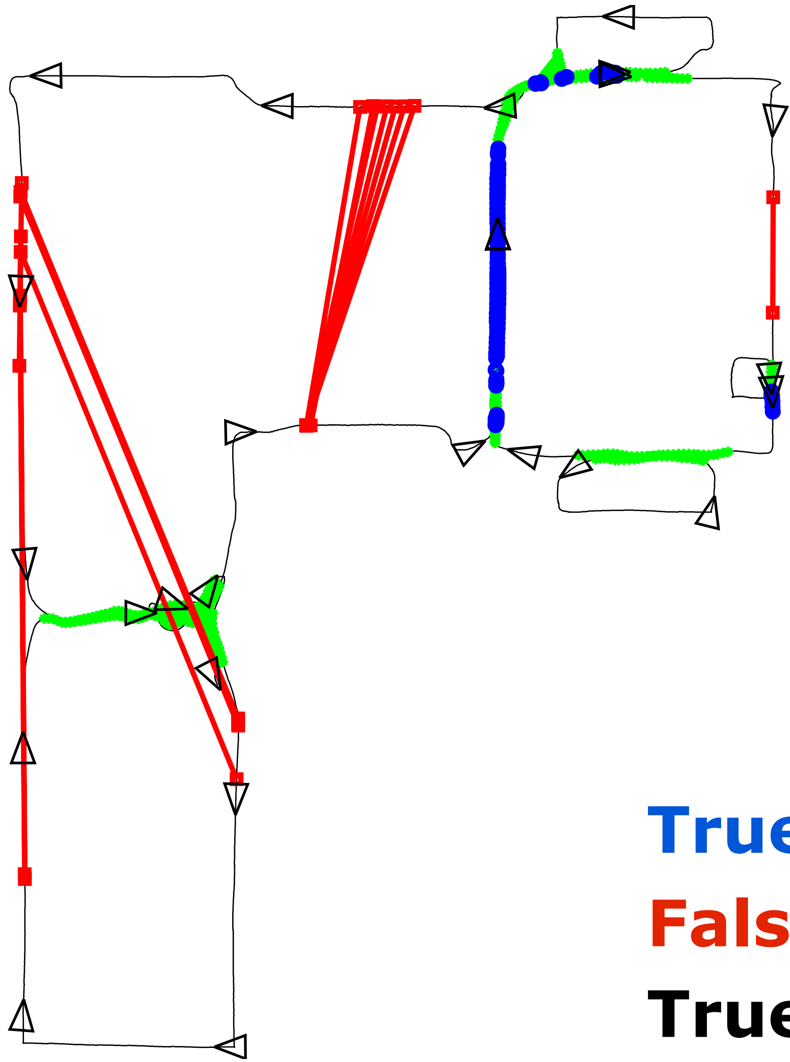
Loop closings

TRINOCULAR indoors

Ground Truth Loop Closings (Indoors)



FAB-MAP



True positives

False positives

True negatives

False negatives

False Positives

Scene 1443

Scene 1244

False Positives

Scene 354

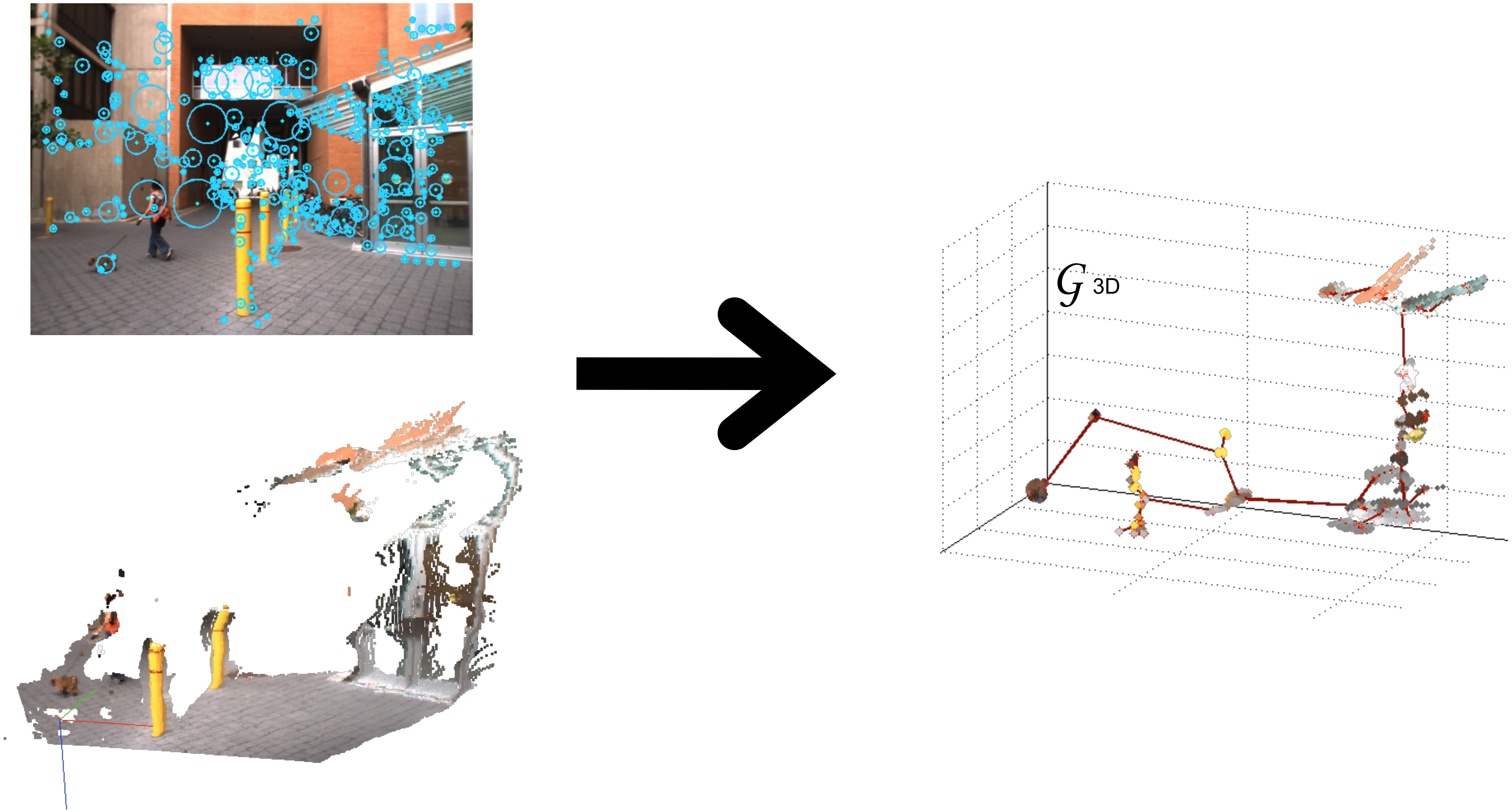
Scene 146

False Positives

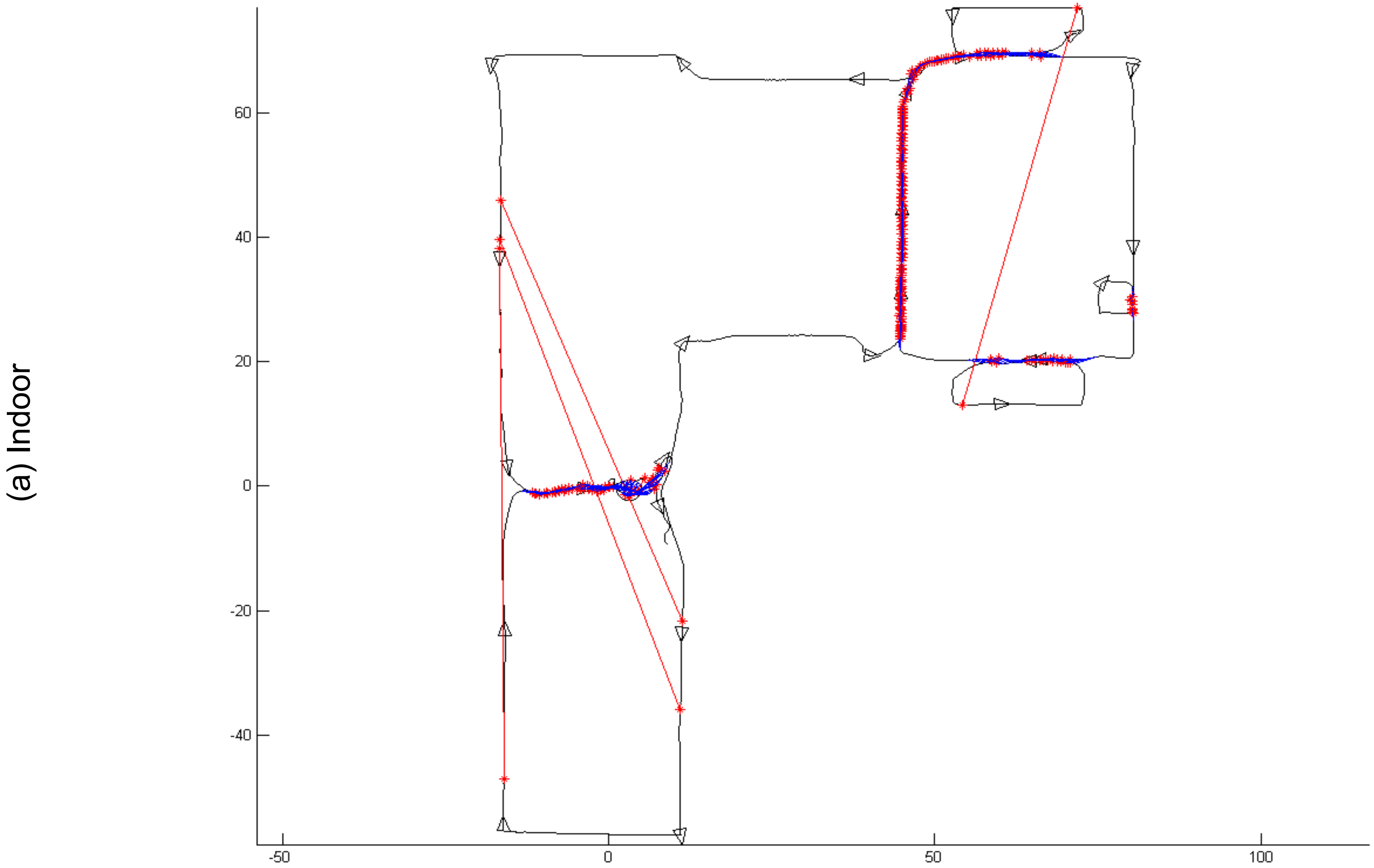
Scene 546

Scene 233

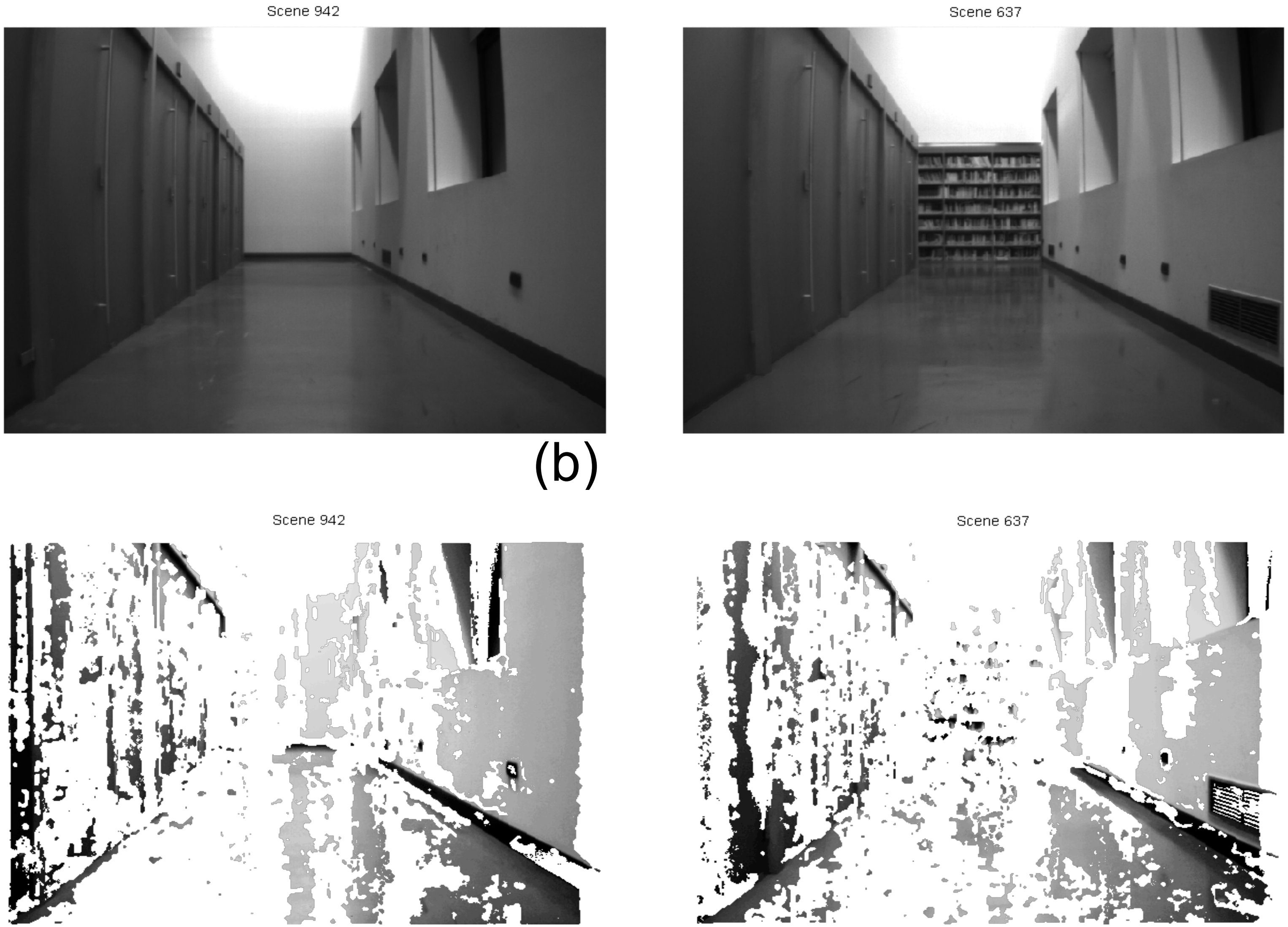
CRF Verification using stereo



CRF 3D: Less (but still) false positives



False positives



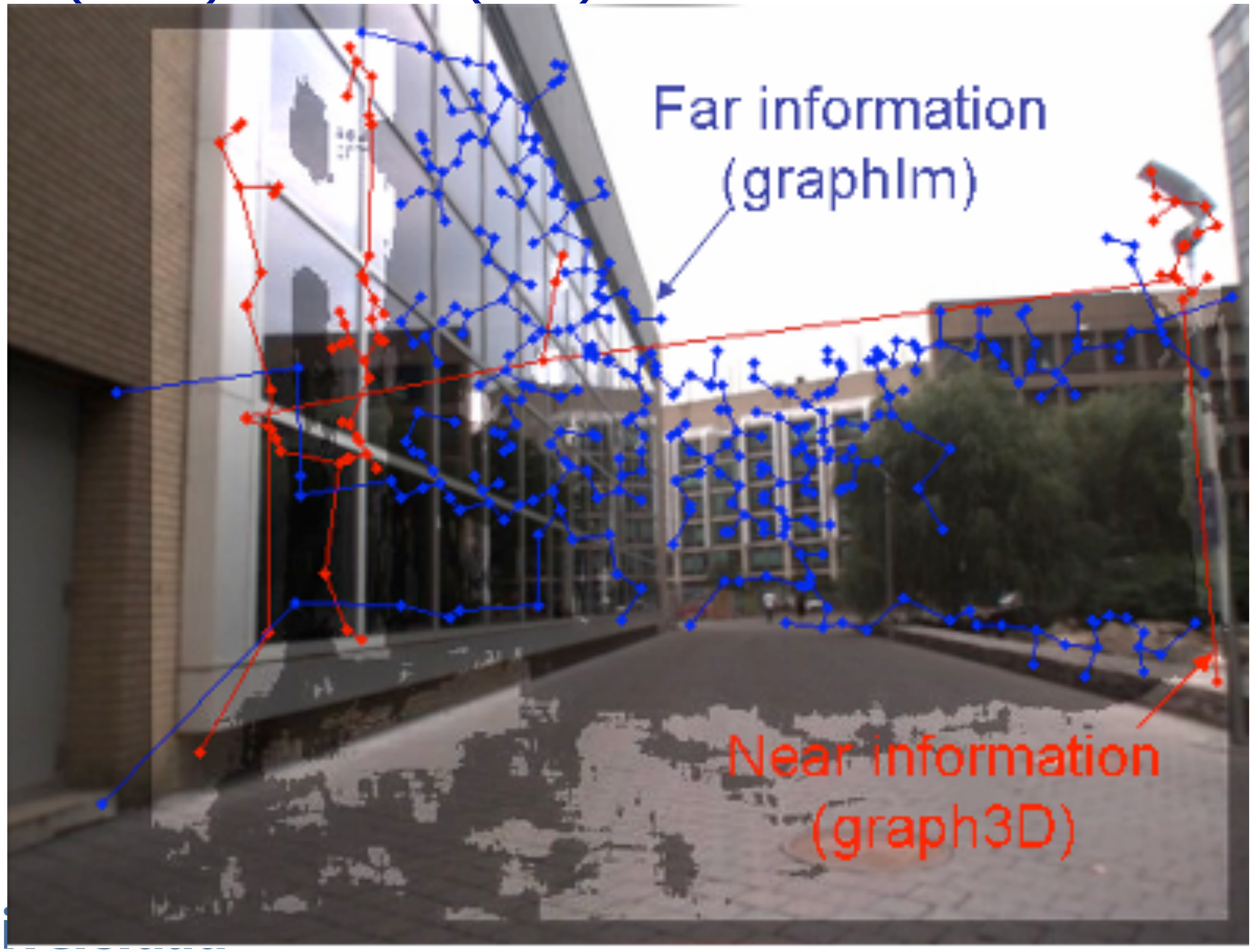
False positives

Scene 292

Scene 219

Using ALL information

- Far (blue) + Near (red)



When does far information help?

Scene 2224

Scene 11

Universidad Zaragoza 1542

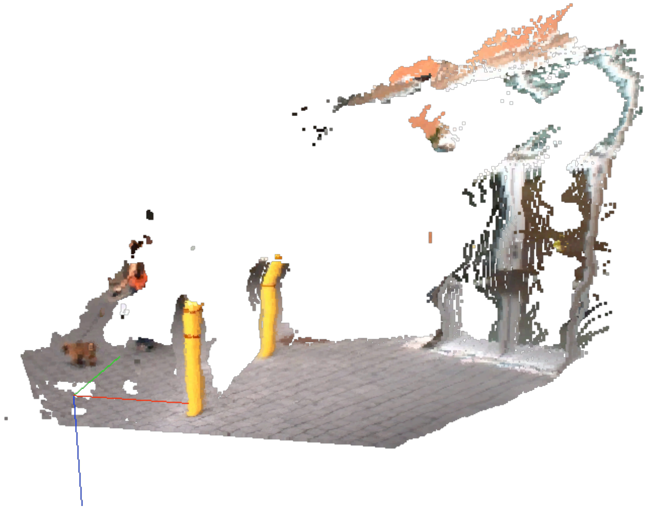
When does far information help?

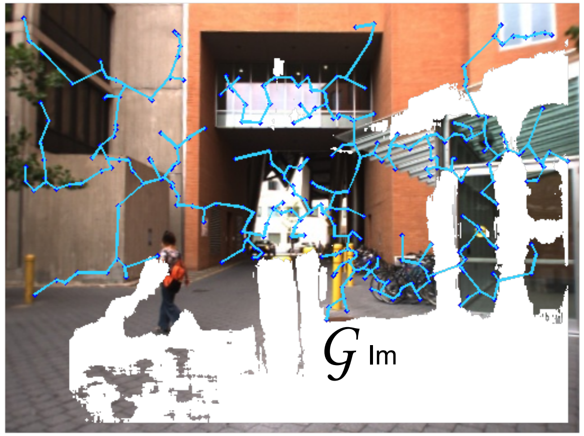
ID 1279649970847876

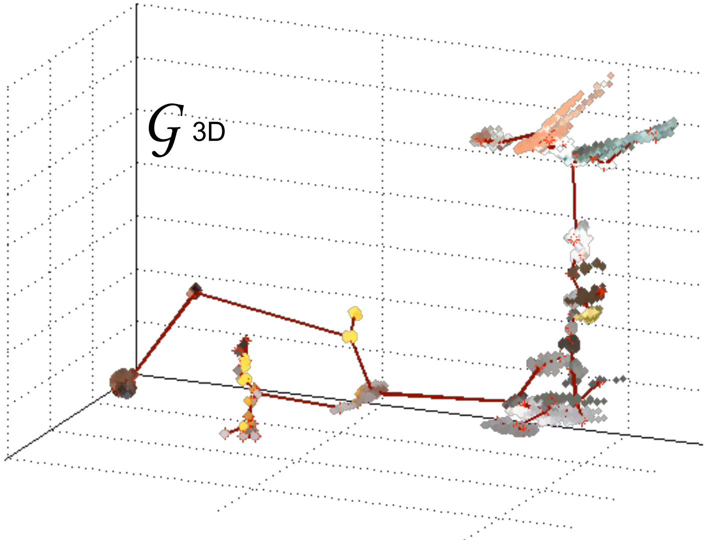
ID 1279556067203571

Universidad Zaragoza 1542

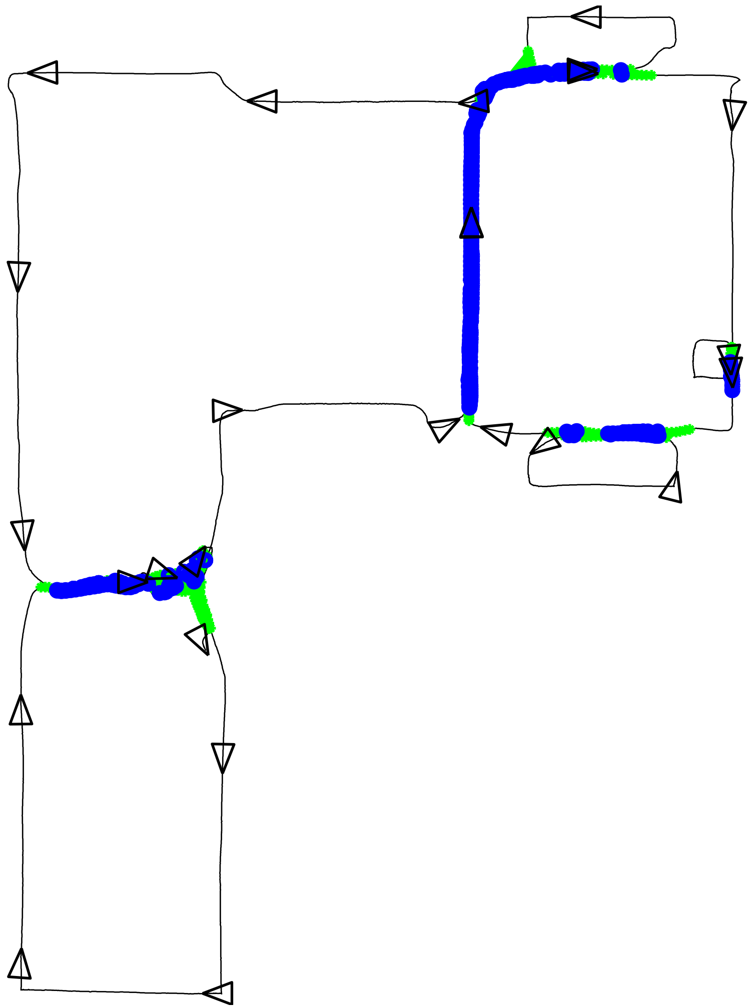
CRF Verification using Image AND stereo





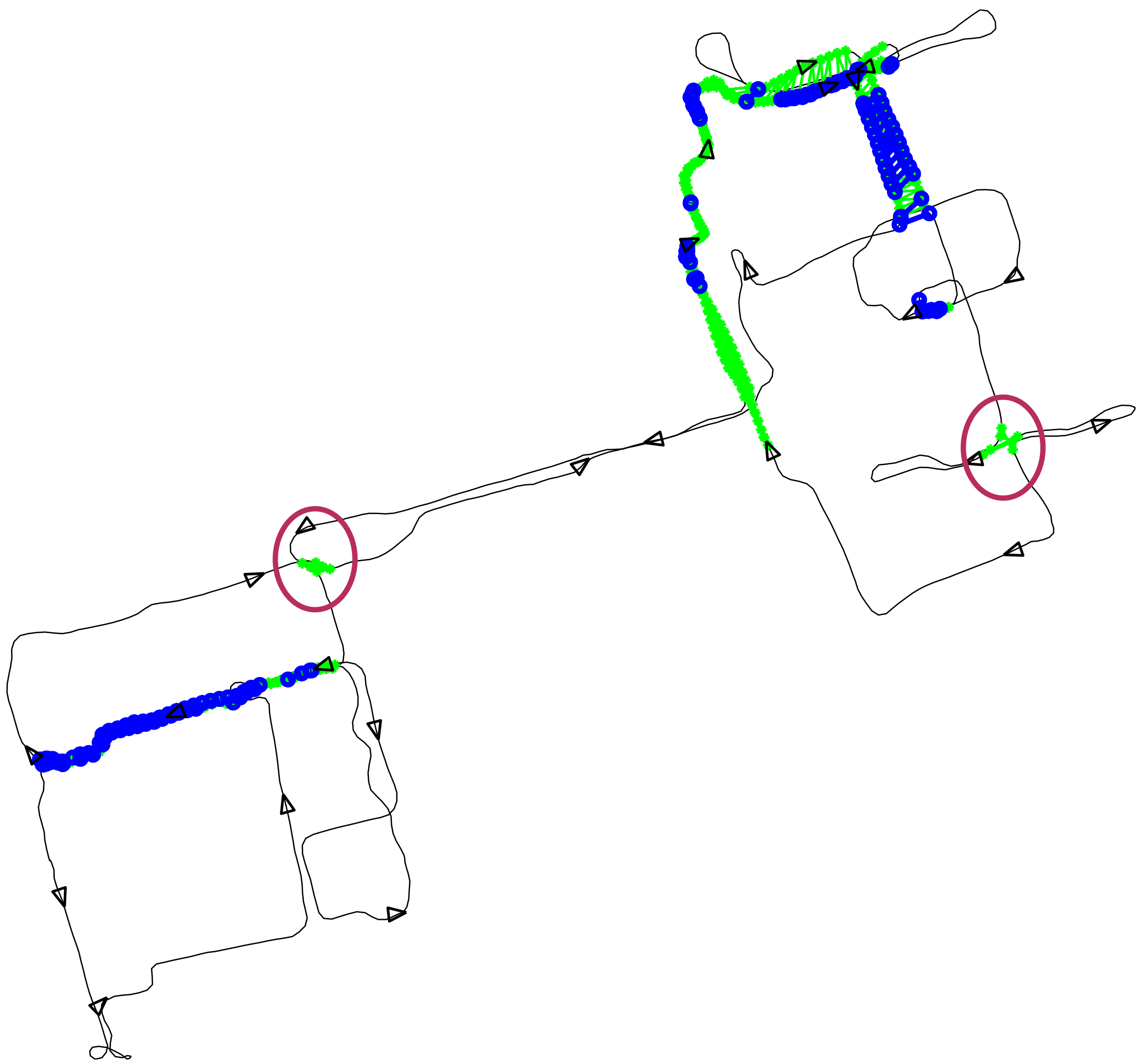


BoW + CRF 3D + IM



No false positives,
All loops detected

BoW + CRF



Missed two loops :-(

Is SLAM solved?

- Editorial by Udo Frese (U. Bremen) with S. Thrun, J. Neira, in Journal **Künstliche Intelligenz 2010**:

- Maybe for indoor static environments, but...

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

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

- Data association is still a challenging problem