Towards a Cognitive Probabilistic Representation of Space for mobile robots

Shrihari Vasudevan, Viet Nguyen & Roland Siegwart

{shrihari.vasudevan,viet.nguyen,roland.siegwart}@epfl.ch

Autonomous Systems Lab, EPFL

CH-1015 Switzerland.

(http://asl.epfl.ch/)

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Vision / Motivation

Cognitive Robot Companions of the future

Navigation

Manipulation

Interaction

Reasoning

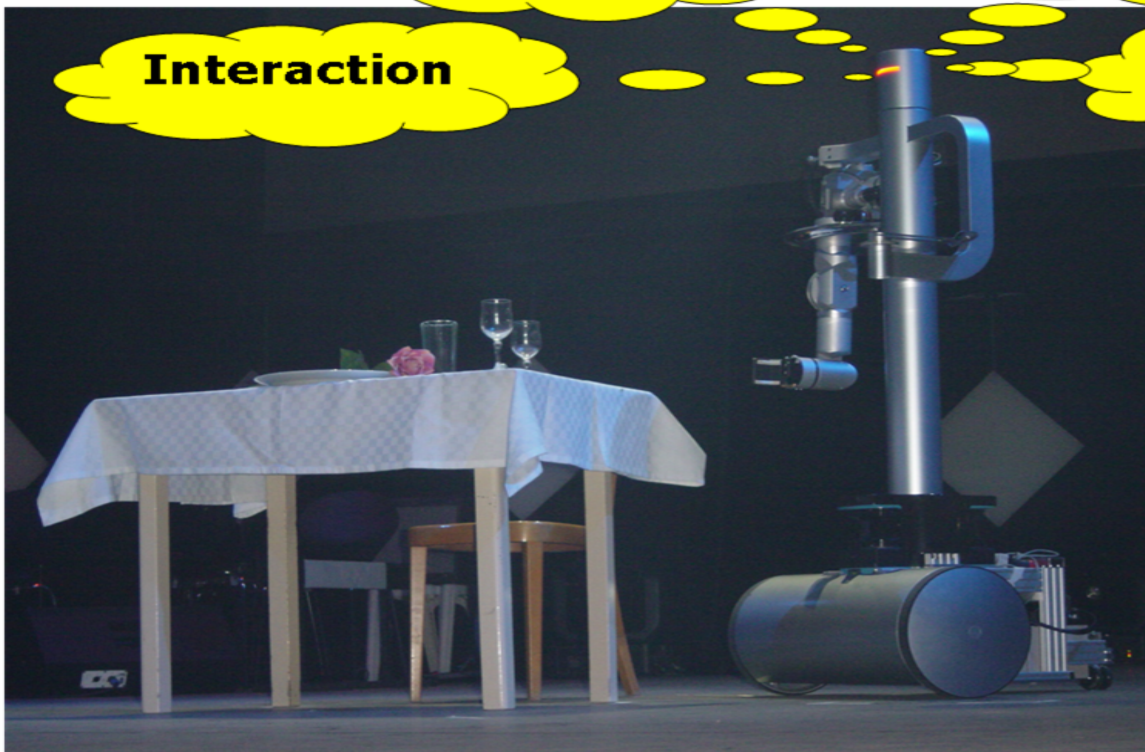
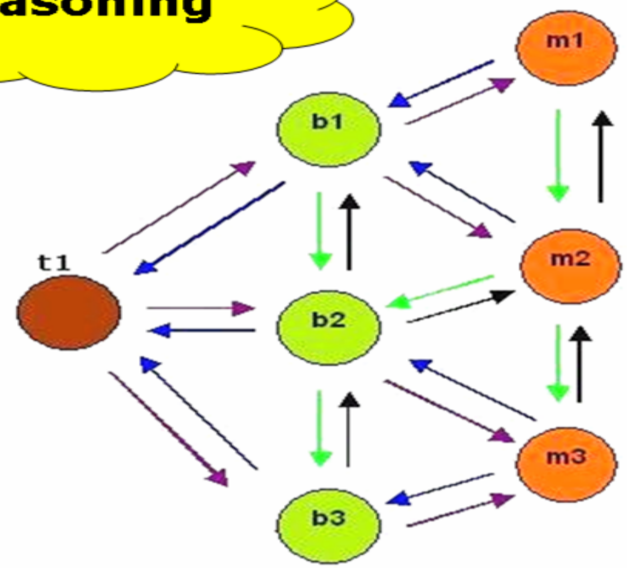


Figure - Robot in a typical home scenario



Spatial relationships between objects, regions & places

Our Work

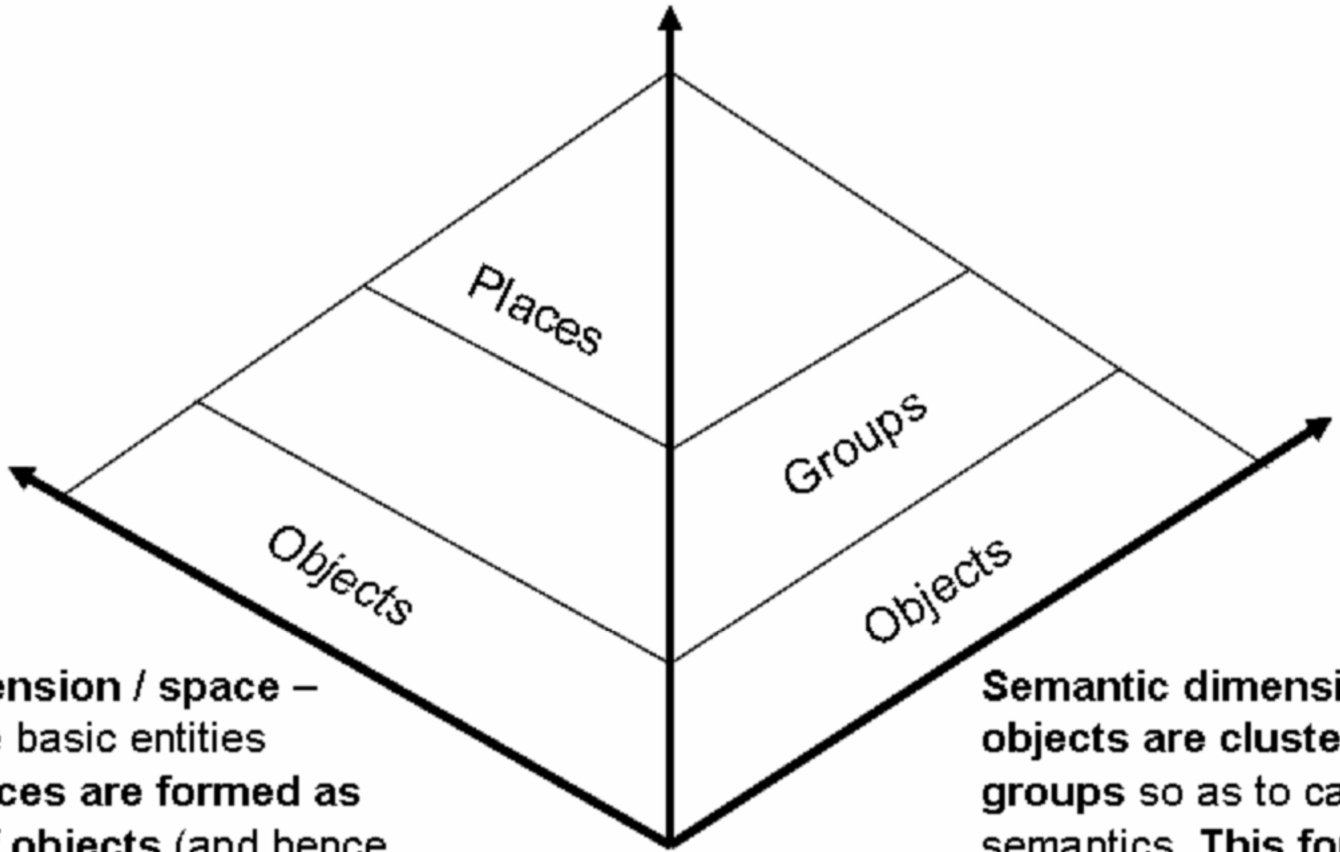
REPRESENTATION FOR ROBOT
Hierarchical Probabilistic Representation of Space

Related Work / Motivation

- Robot mapping – well researched problem !
  - Metric Maps [Arras03, Martinelli04]
  - Topological maps [Choset01, Tapus05]
  - Hybrid maps [Thrun98, Tomatis03]

- More recent trends:
  - Hierarchical representations [Kuipers00, Galindo05]
  - "High level feature" extraction [Lowe04, Anguelov04]
    - (object recognition / classification ; door detection etc.)
  - Scene Interpretation (semantics) [Torralba03, Stachniss05]
  - Cognitive Maps [Kuipers83, Yeap01]

Approach

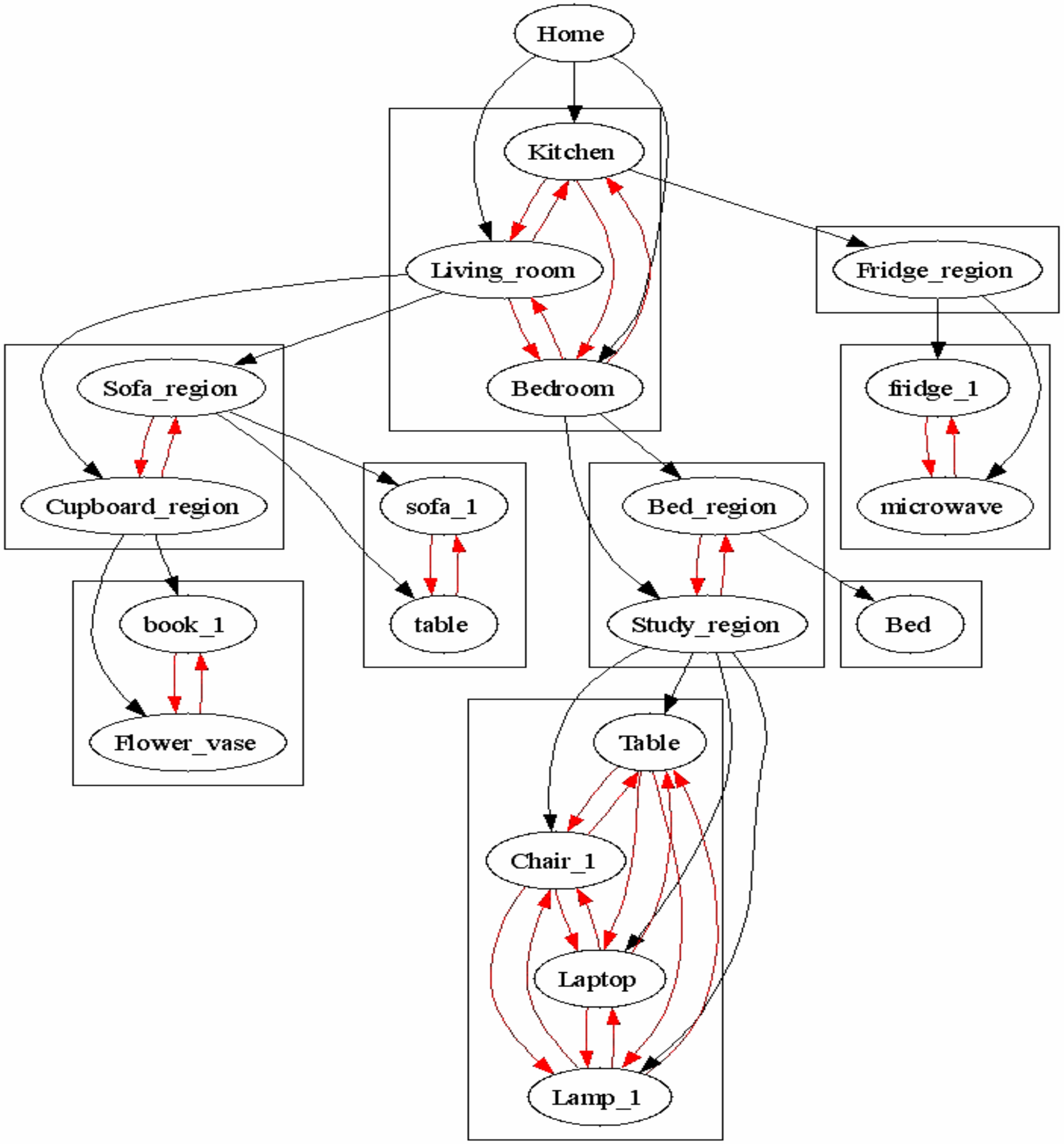


Physical dimension / space – objects are the basic entities (features). **Places are formed as collections of objects** (and hence groups) . **This formation is primarily PERCEPTUAL.**

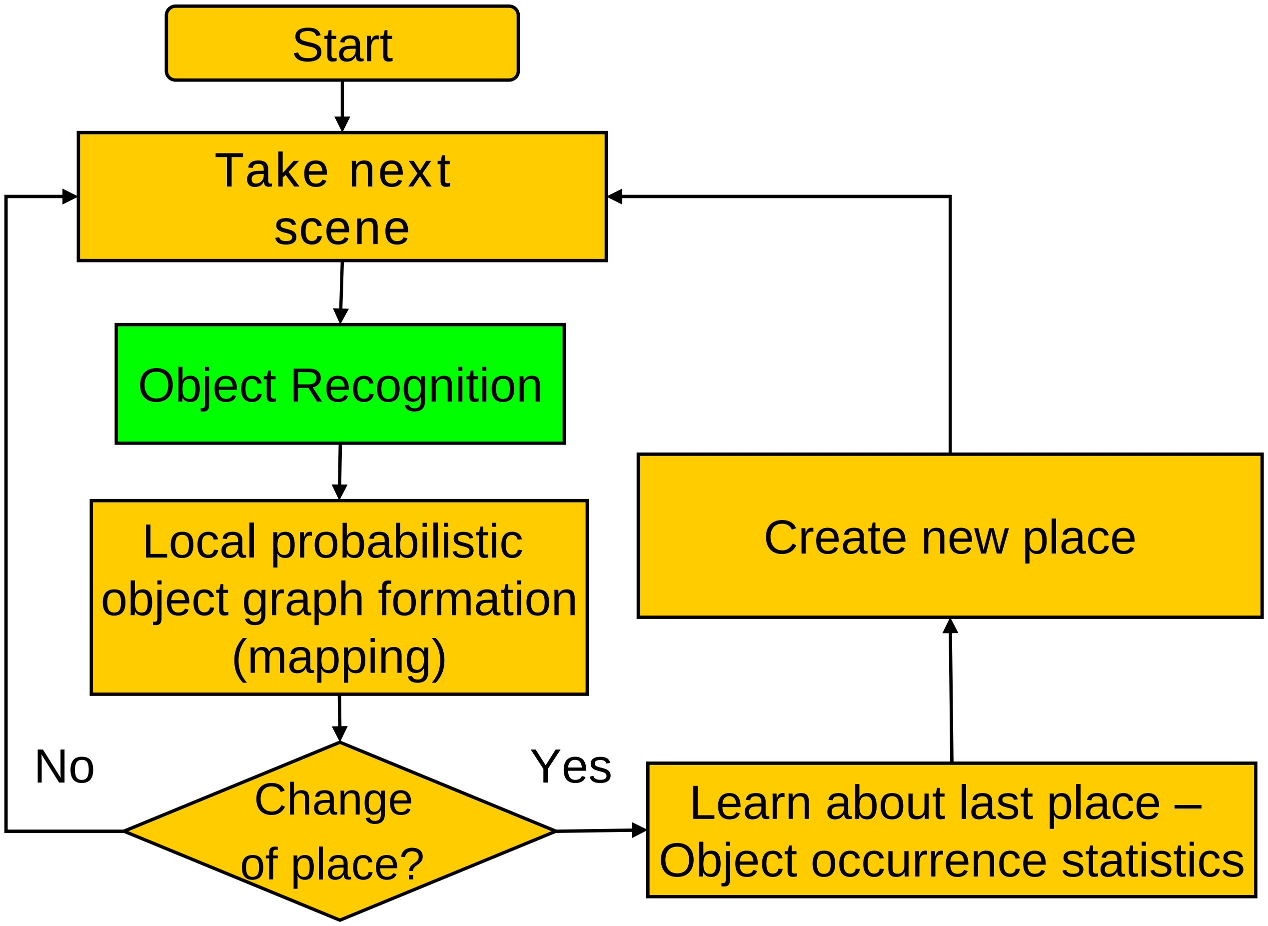
Semantic dimension / space – objects are clustered into groups so as to capture their semantics. **This formation is primarily SEMANTIC.**

Both dimensions are represented in the same hierarchical representation. Only, the different levels of abstraction are formed using different metrics – which may be perceptual (such as visibility / distance) or semantic (similarity – property / functional ; distance [nearness] / touch)…

Approach – example



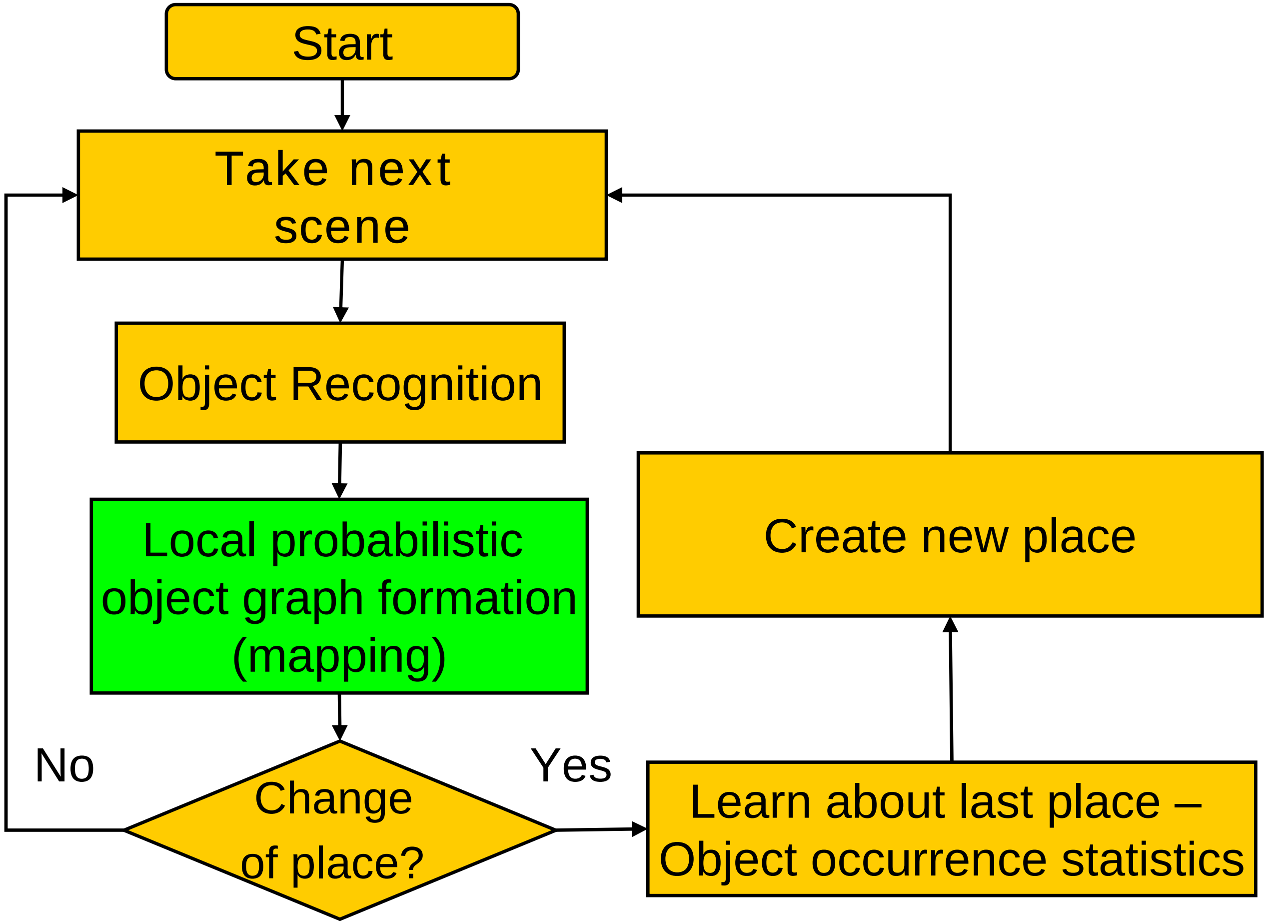
Strategy – mapping



Object Recognition

- Using Lowe's SIFT features
- Using a limited set of textured objects
  - Different cartons / chair / mug / shelf / table & book.
- SIFT (Lowe2004)
  - "Local features" approach
  - Does NOT learn general properties of objects
  - Transforms a naively obtained set of features into a robust feature set that incorporates invariance to scale / rotation and to some extent deals with illumination changes and changes in viewing direction.
  - Our experience – good for textured objects and specifically for object detection.

Strategy – mapping



Probabilistic graphical representation

- Representation encodes objects and relationships between them.
- Relationships are relative spatial information implemented as distance and angles (in 3D).
- Both existential (discrete belief measure) and precision beliefs (classical representation - mean / covariance matrix) are maintained and treated separately.

Representation – some details

- Odometry model – Standard differential drive model.
- Stereo model – suggested by Jung and Lacroix for their work on SLAM using stereo vision.
- Belief representation of objects in local maps –

$$X_o = f(X_c)\, where\; f = M_{RO} * M_{CR} \quad \text{(object in local place reference)}$$

$$P_o = F_1 P_1 F_1' + F_2 P_2 F_2' \quad \text{(covariance matrix – uncertainty in object position)}$$

where

$$F_1 = J_{X_1}(f) \quad \& \quad F_2 = J_{X_2}(f)$$

where

$X_1 \rightarrow (X_R, Y_R, \theta_R)$ is the robot pose

$X_2 \rightarrow (X_C, Y_C, \theta_C)$ is the object position in the camera reference frame

P_1 is the covariance matrix representing the uncertainty in robot position.

P_2 is the covariance matrix representing the uncertainty in the object position.

Representation – more details

- Belief representation of relative spatial information between objects in local map

$$X_1(x_1, y_1, z_1) \;\&\; X_2(x_2, y_2, z_2) \text{ represent two objects}$$

$$f(X_1, X_2) \rightarrow \text{relative spatial information between the two objects.}$$

$$P_1 \;\&\; P_2 \rightarrow \text{uncertainty in object positions}$$

$$Bel_1(f) = F_1 P_1 F_1' + F_2 P_2 F_2'$$

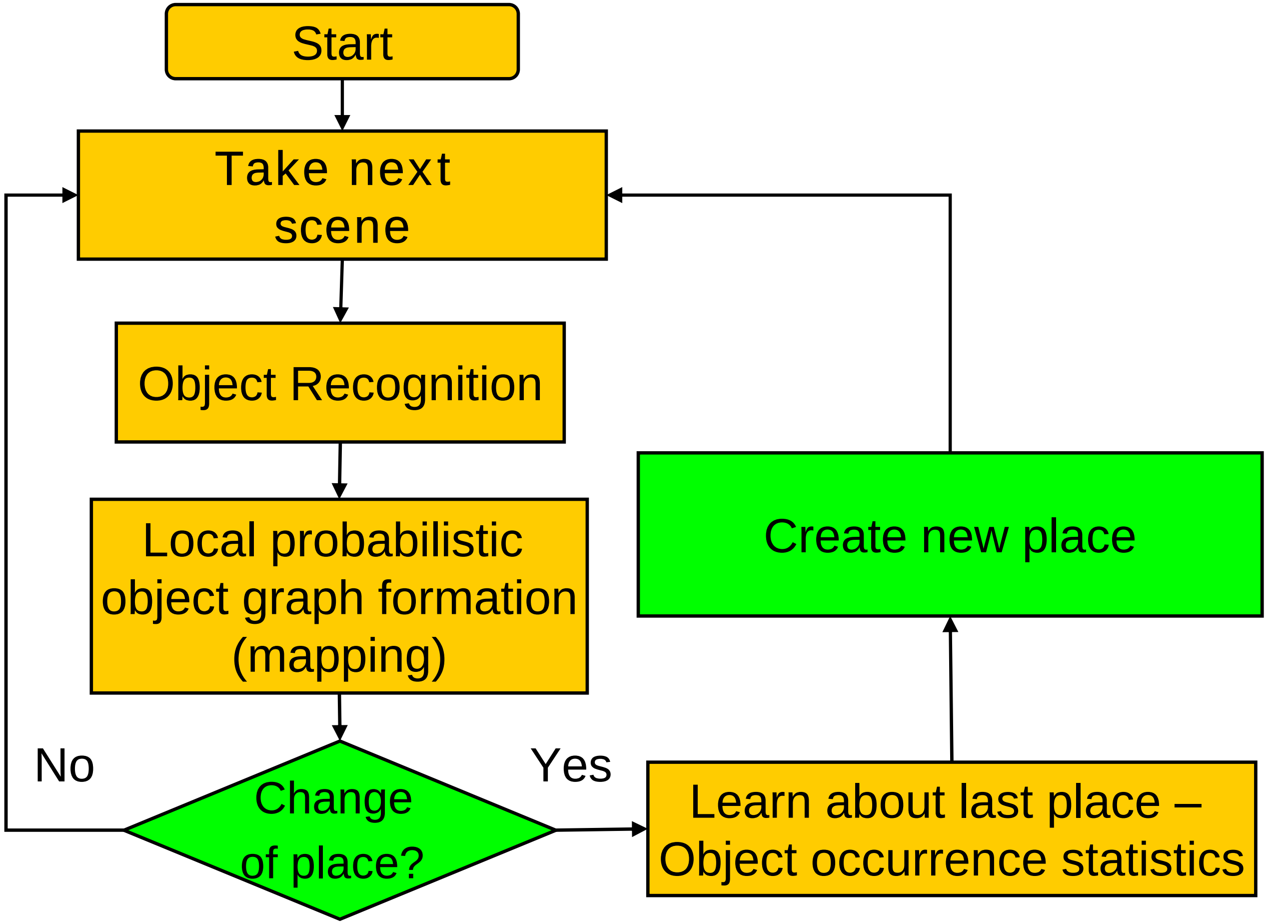
$$(precision)$$

$$where\; F_1 = J_{X_1}(f) \;\&\; F_2 = J_{X_2}(f)$$

$$and \quad Bel_2(f) = \min(belief\ in\ existence\ of\ objects)$$

$$(existence) \qquad (from\ OR\ system)$$

Strategy – mapping



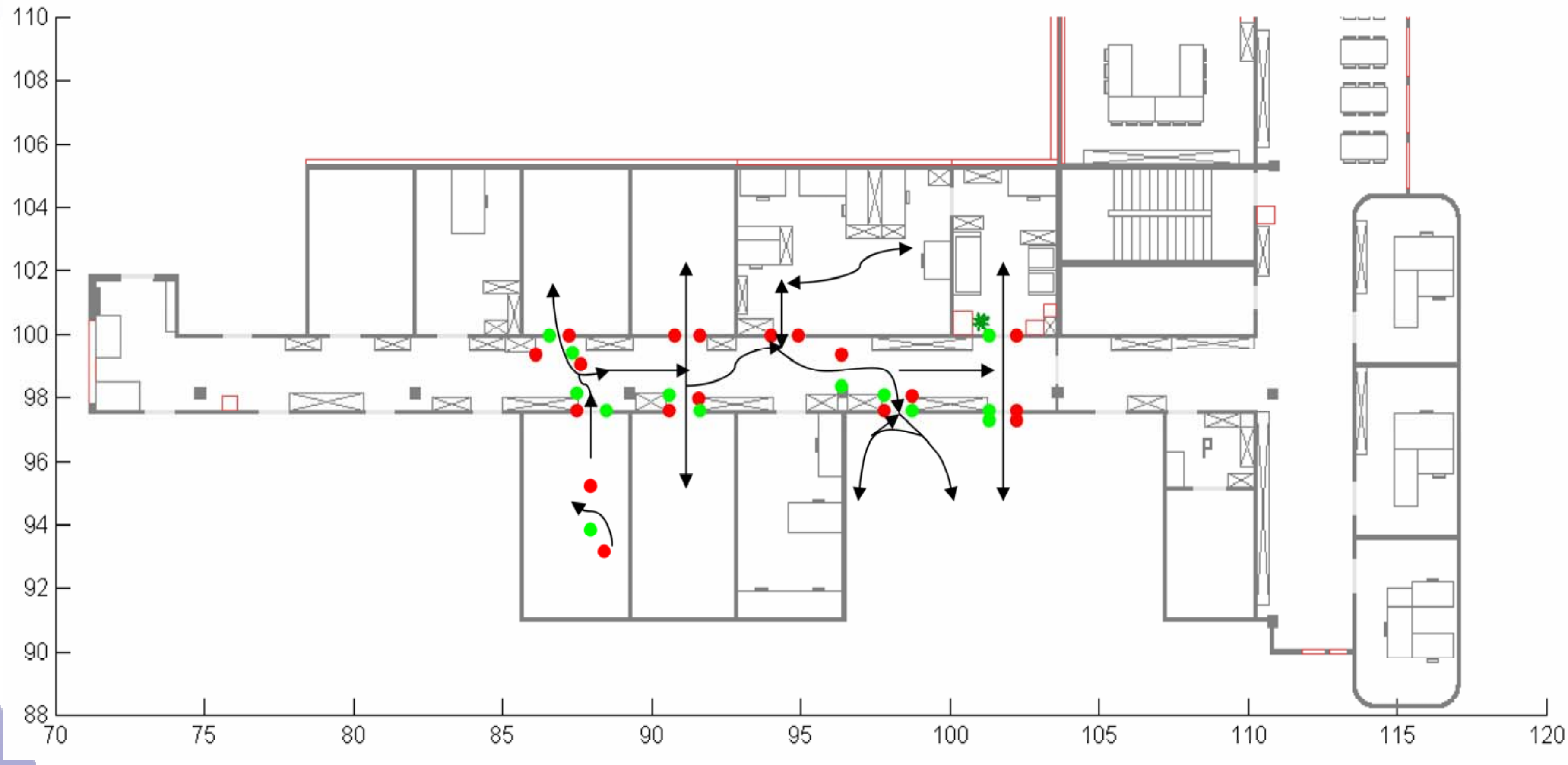
Place formation

- What is a place ?
  - Distance
  - Low level (SIFT) Feature based place formation (transition detection)
  - Door detection

- Current Scheme
  - Door detection ; Doors are incorporated into the representation only when crossing it to another ‘place’.

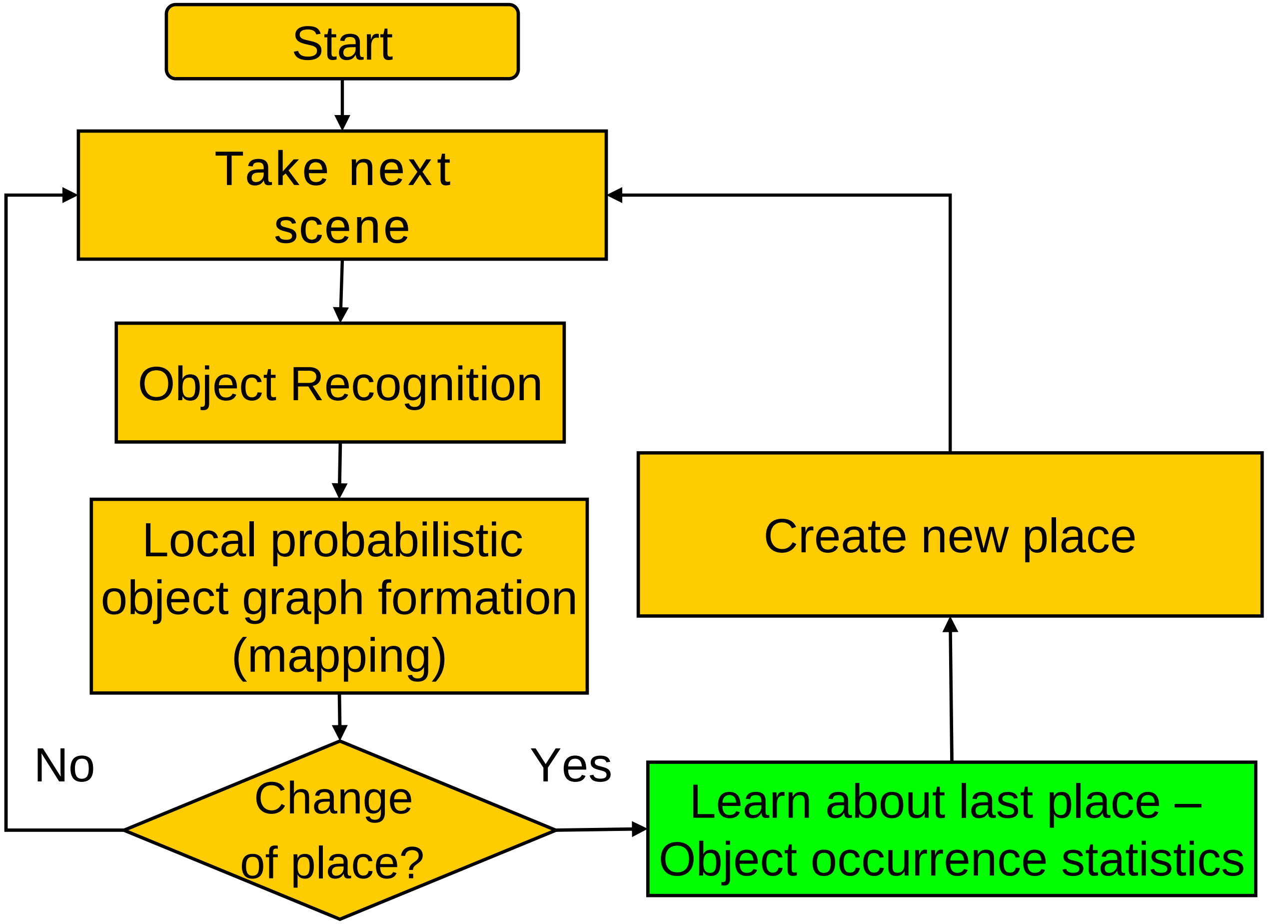
- Door detection
  - Line extraction / laser scanner based.
  - Method (in brief)
    - Detect lines (recursive splitting method)
    - Apply heuristics (identify doorways)
    - Track door hypotheses (likelihood based false positive elimination)
    - Obtain doors when crossing.

Door detection



A pair of colored dots / circles constitutes a door. The red ones are the references of the place that the robot explores on crossing the door.

Strategy – mapping



Concept Learning – understanding places

- Occurrence statistics of objects are stored as a means of understanding a place in terms of its objects.

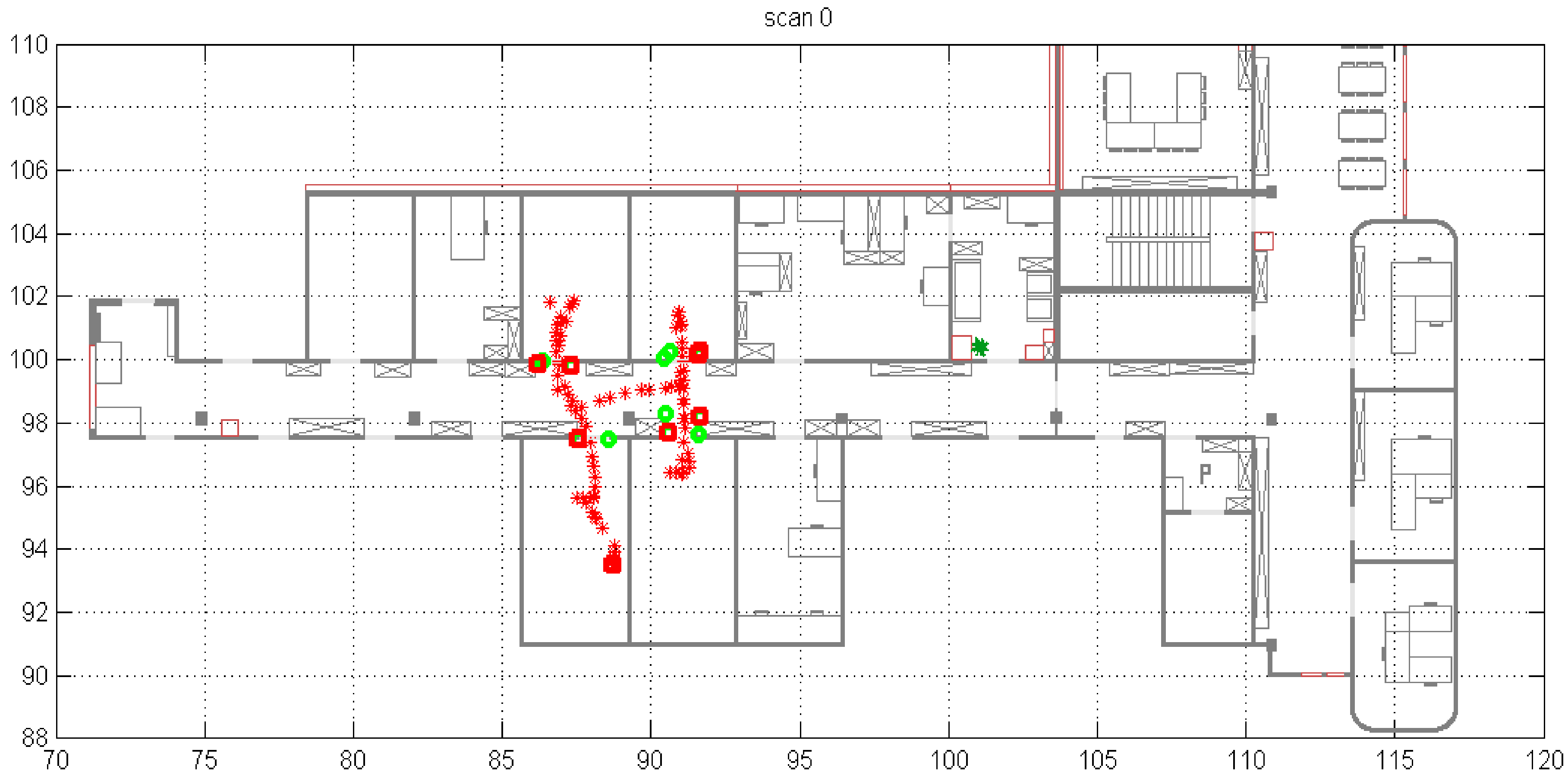
- Useful for future exploration / classification of places / reasoning about places...

- Laplace succession law is used to compute likelihoods of different places given the occurrence statistics of different objects.

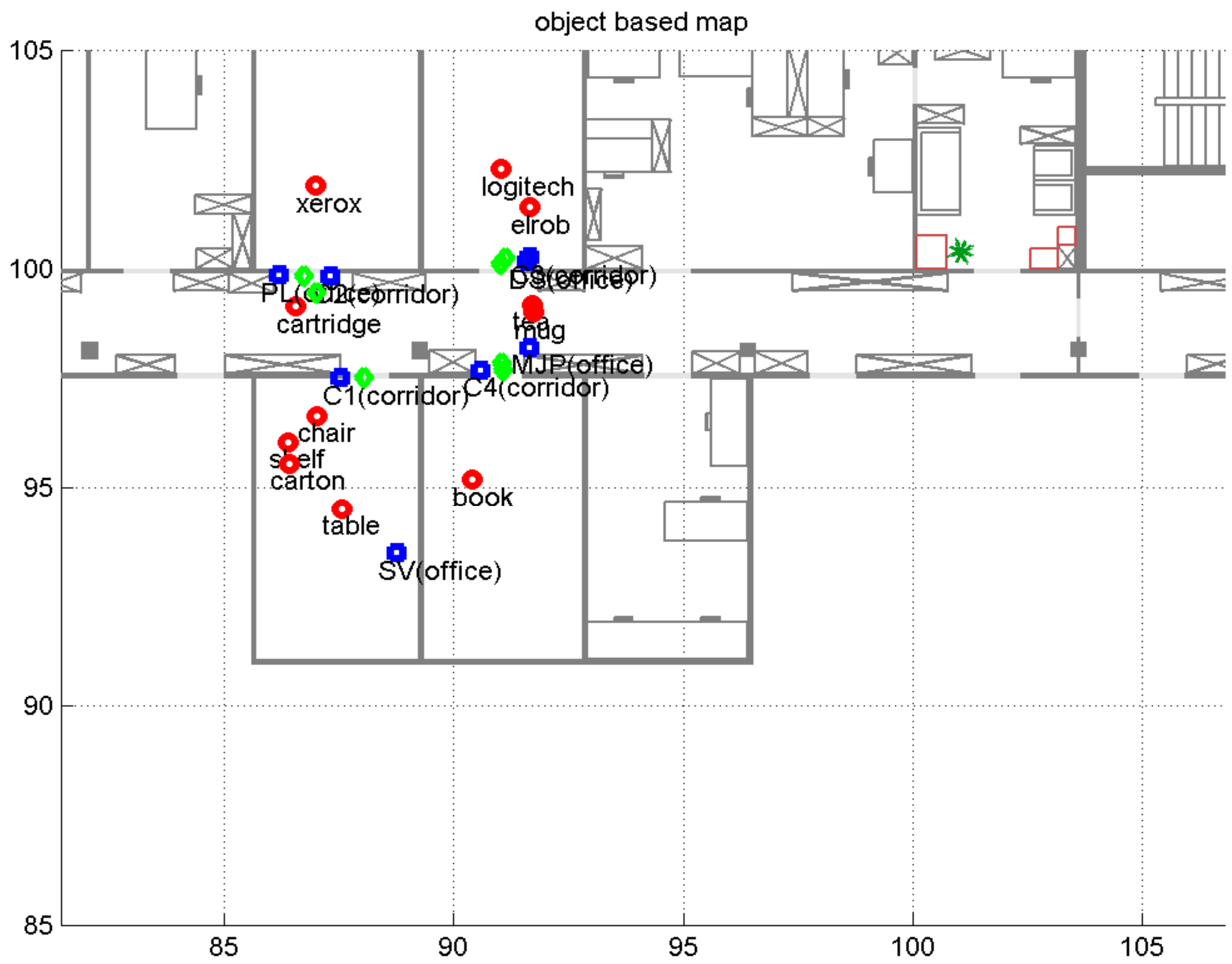
$$P(place|object) = \frac{N(object)+1}{N(place)+2}$$

$$(Likelihood\ of\ place)$$

Mapping – results (robot path)

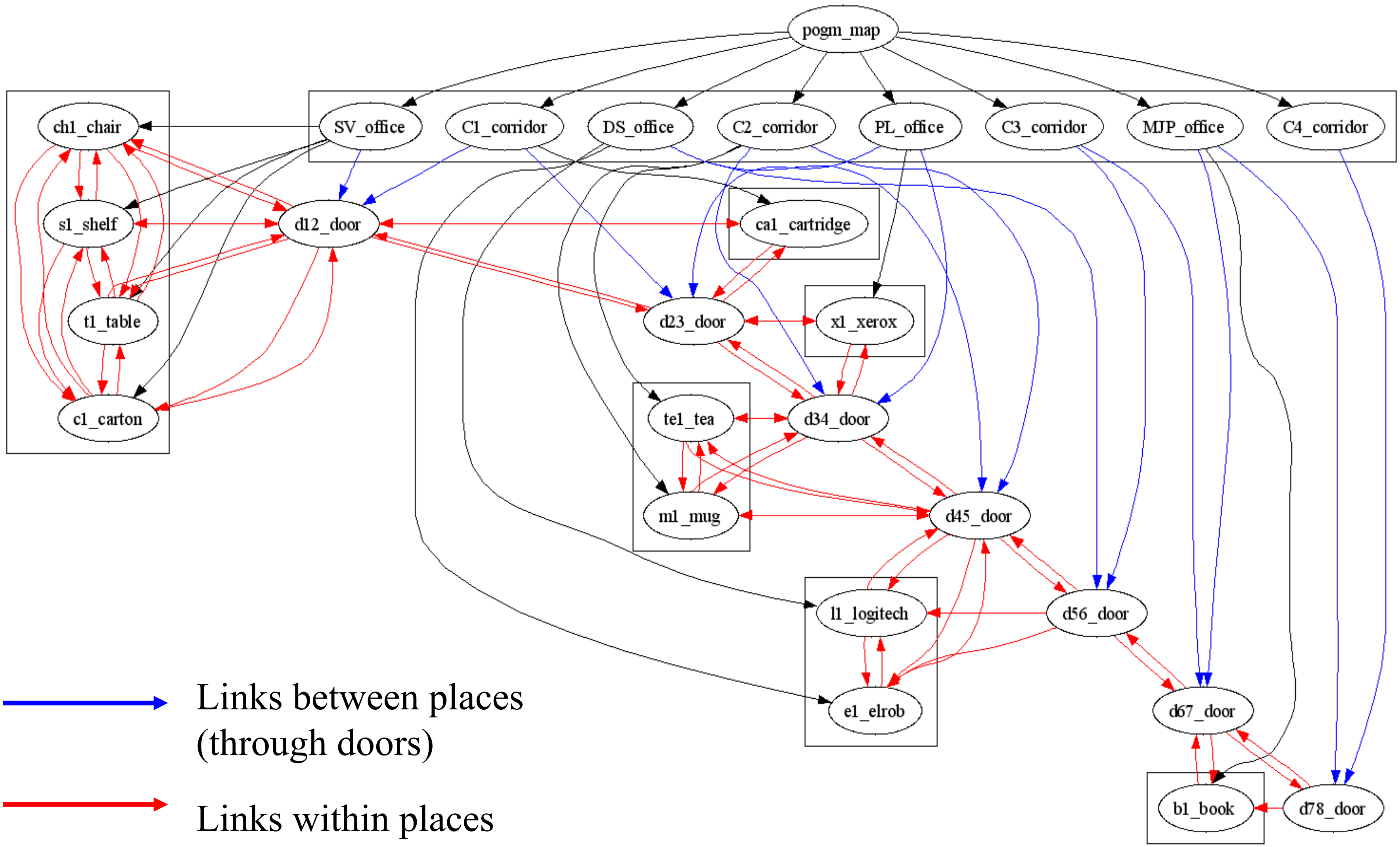


Mapping results (object based map)



- object
- place reference
- door

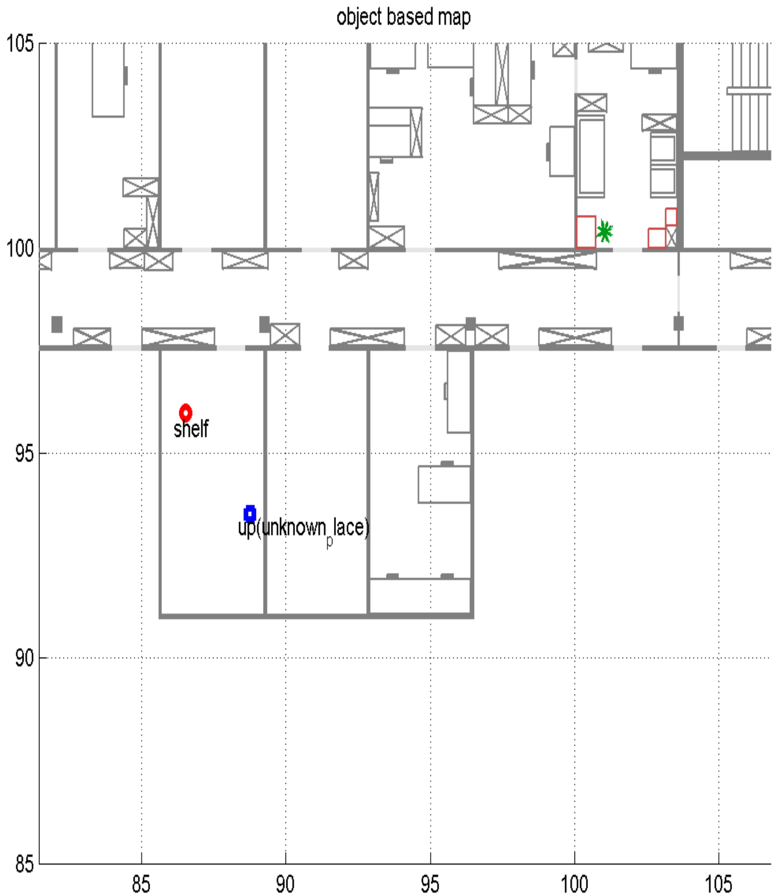
Mapping results (pogm representation)



Links between places (through doors)

Links within places

Scene 8

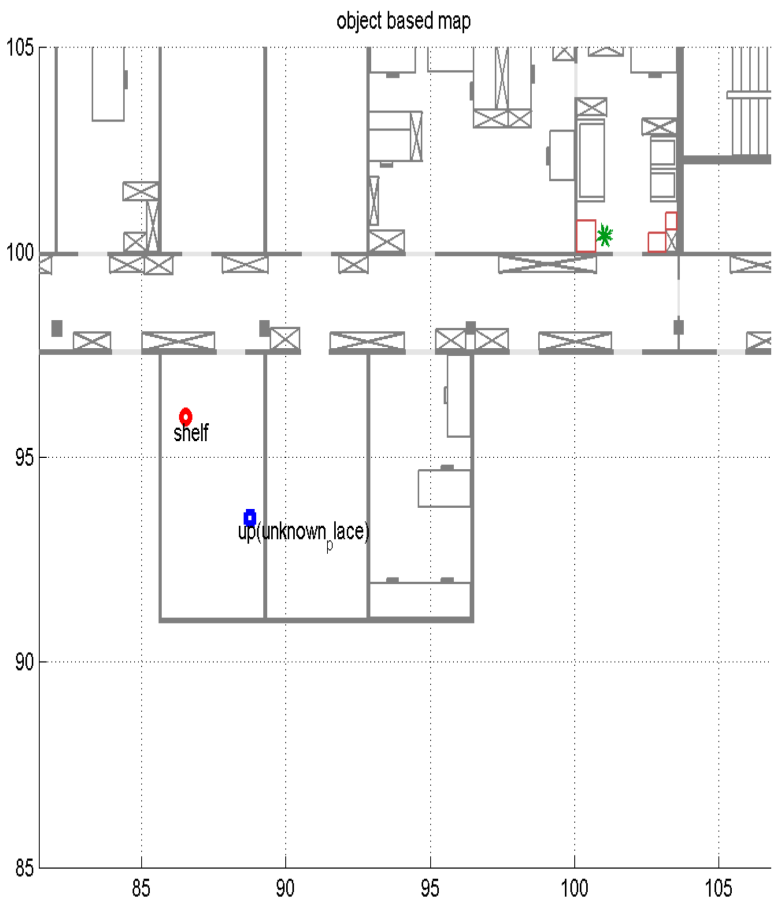


object

place reference

door

Scene 9

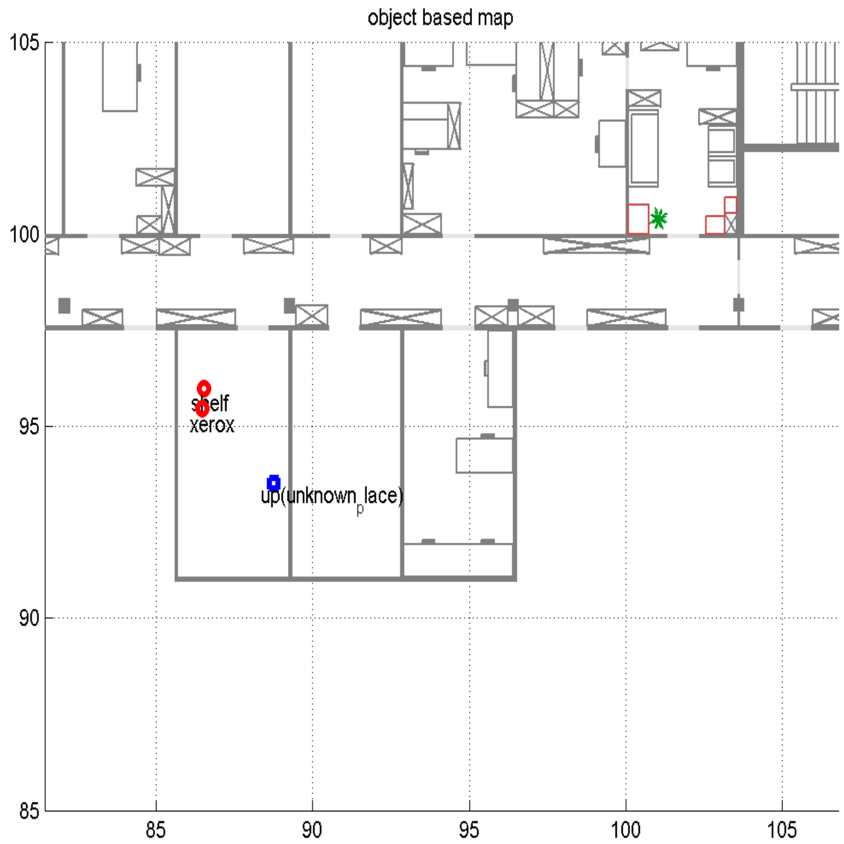


object

place reference

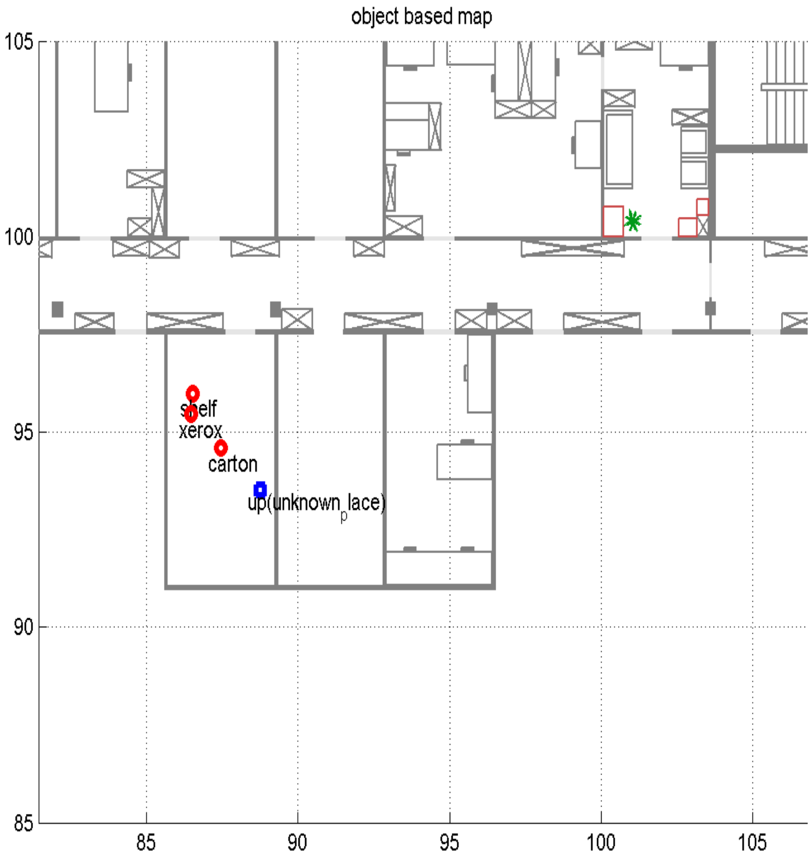
door

Scene 10



Unknown place is classified as an office with a belief of 0.74

Scene 11



object place reference door

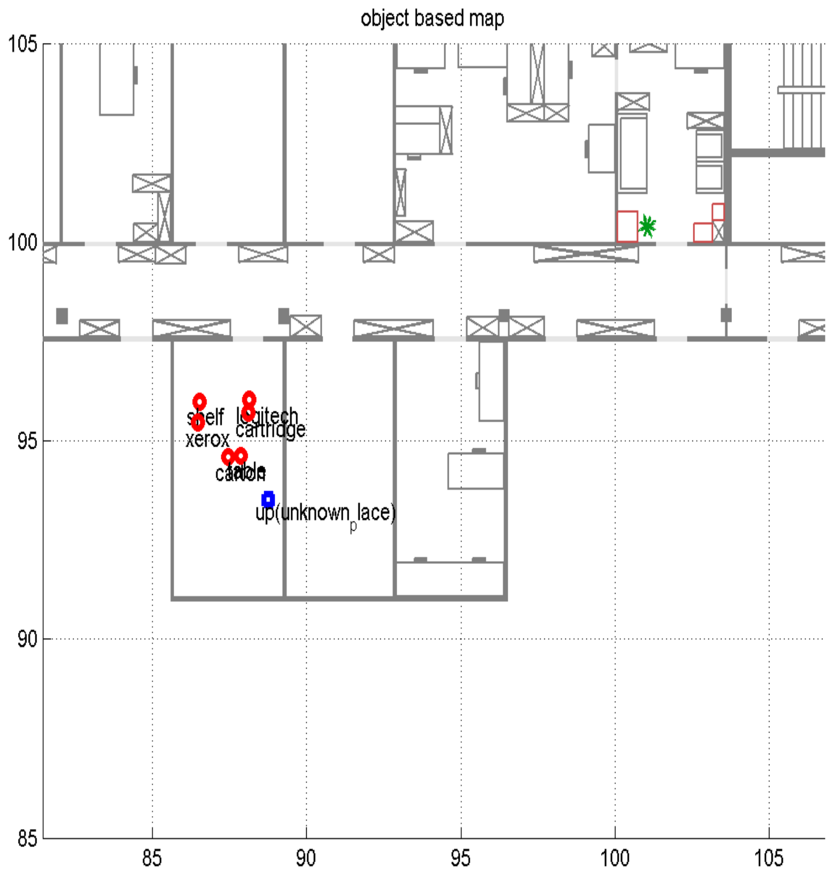
Scene 12

object

place reference

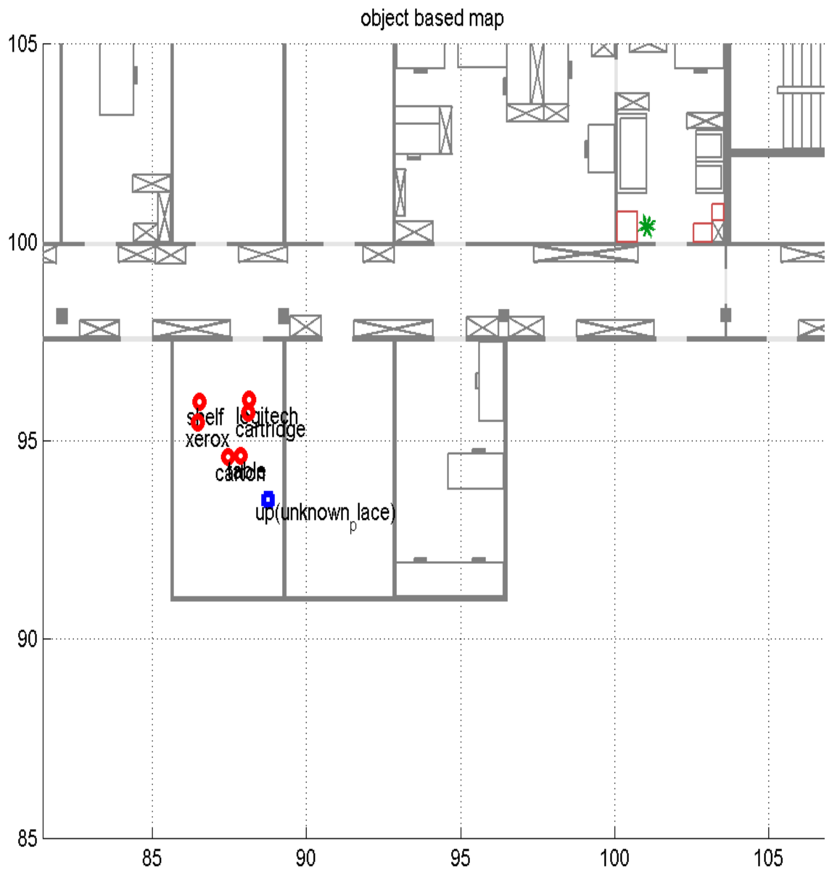
door

Scene 13



object   place reference   door

Scene 14

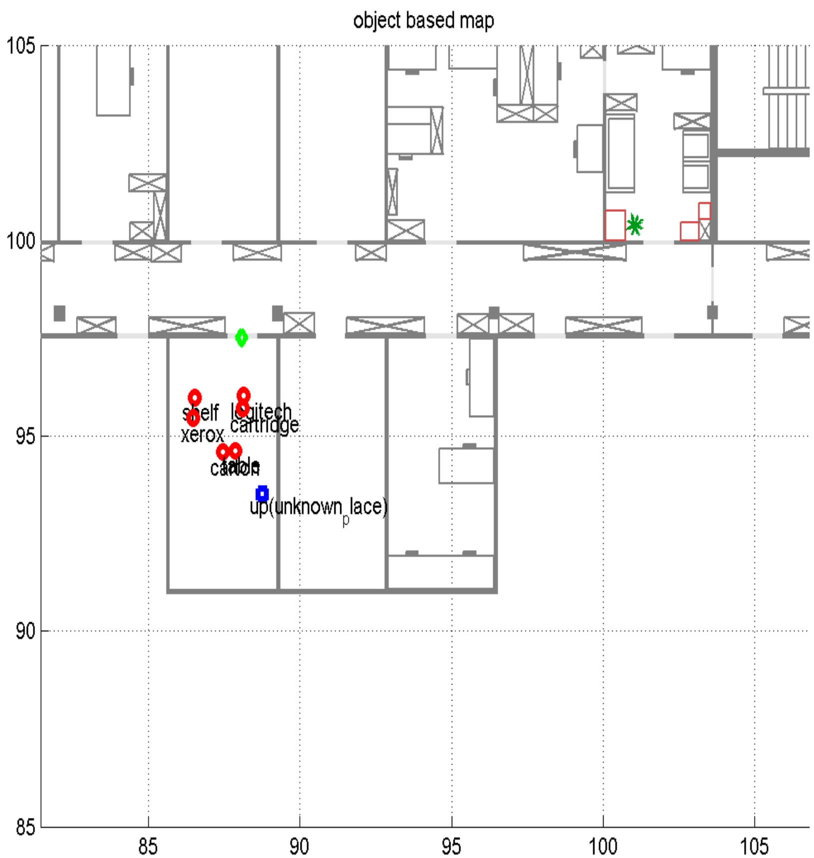


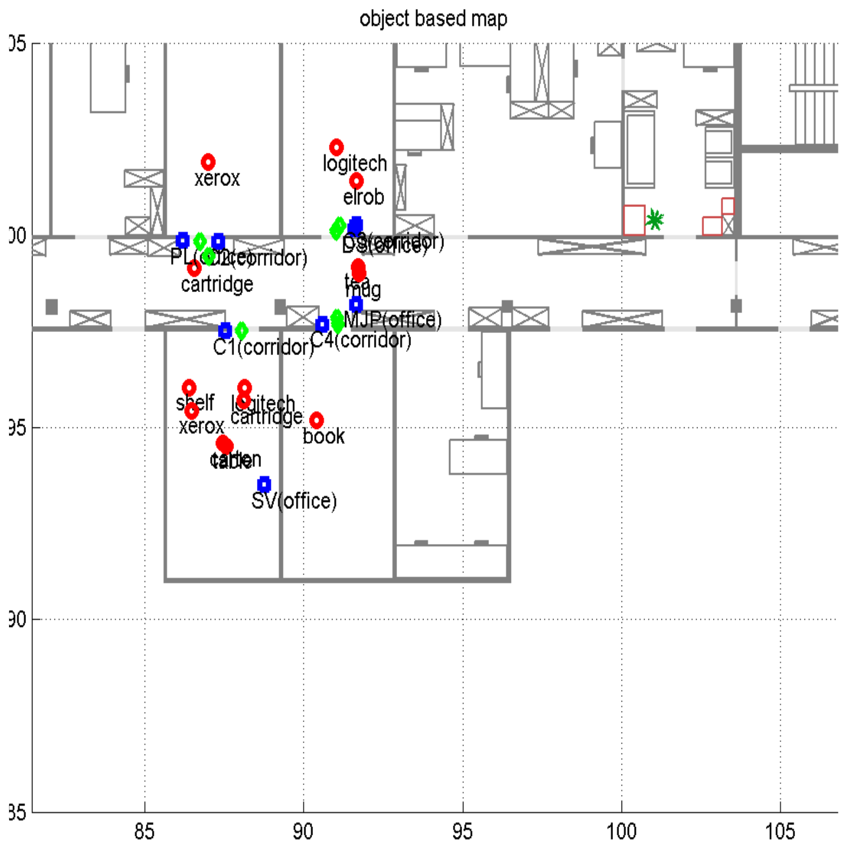
object

place reference

door

Scene 20

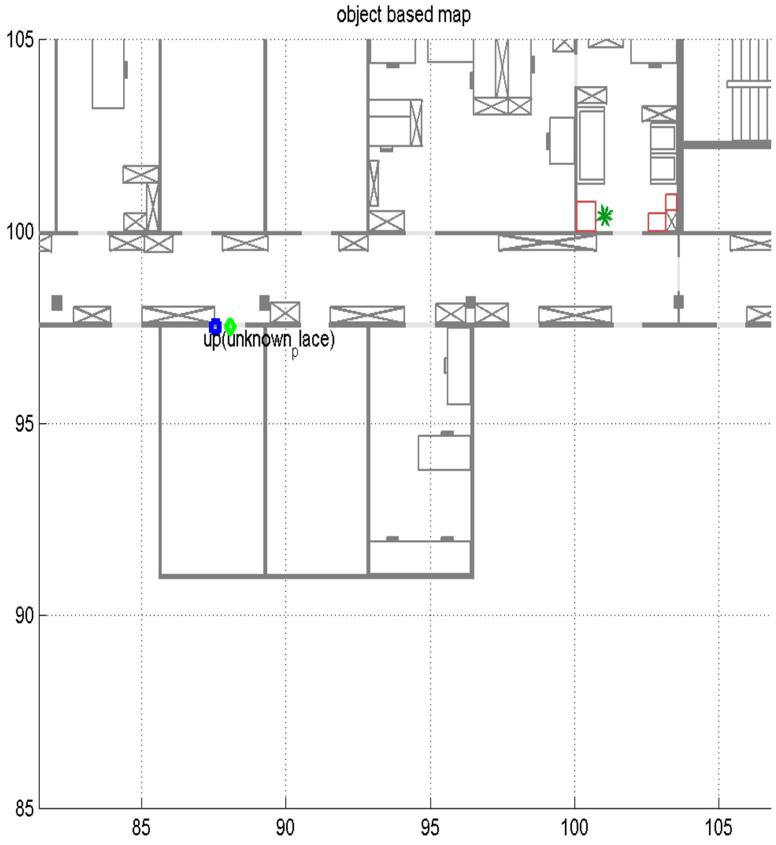


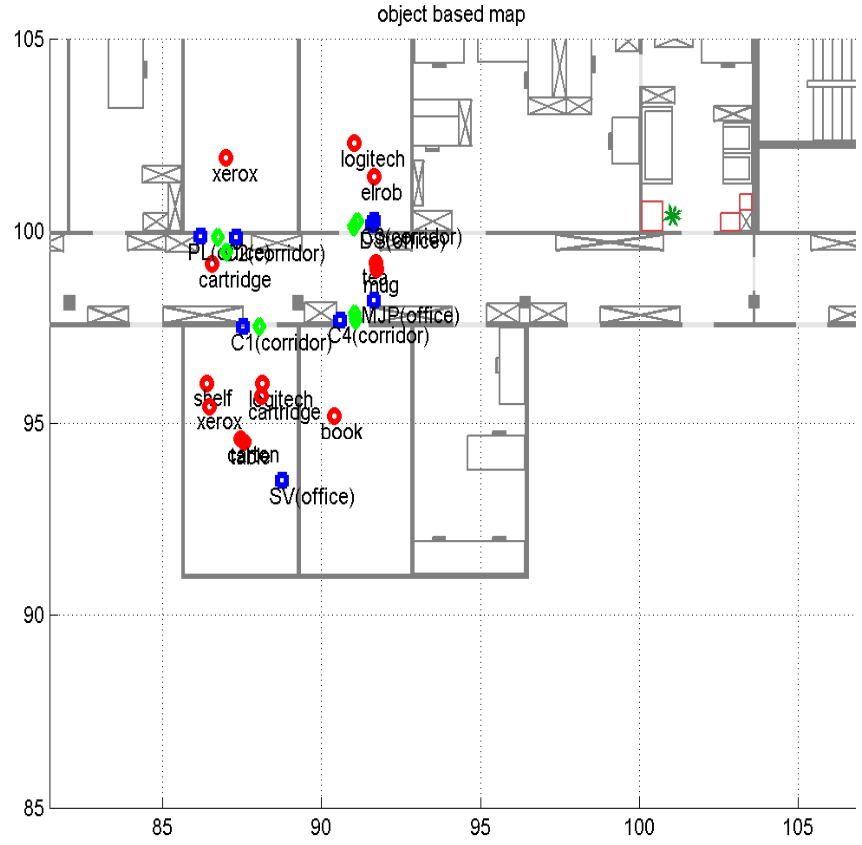


Unknown place is recognized as SV (office) with a belief of 0.60.

Map update is performed.

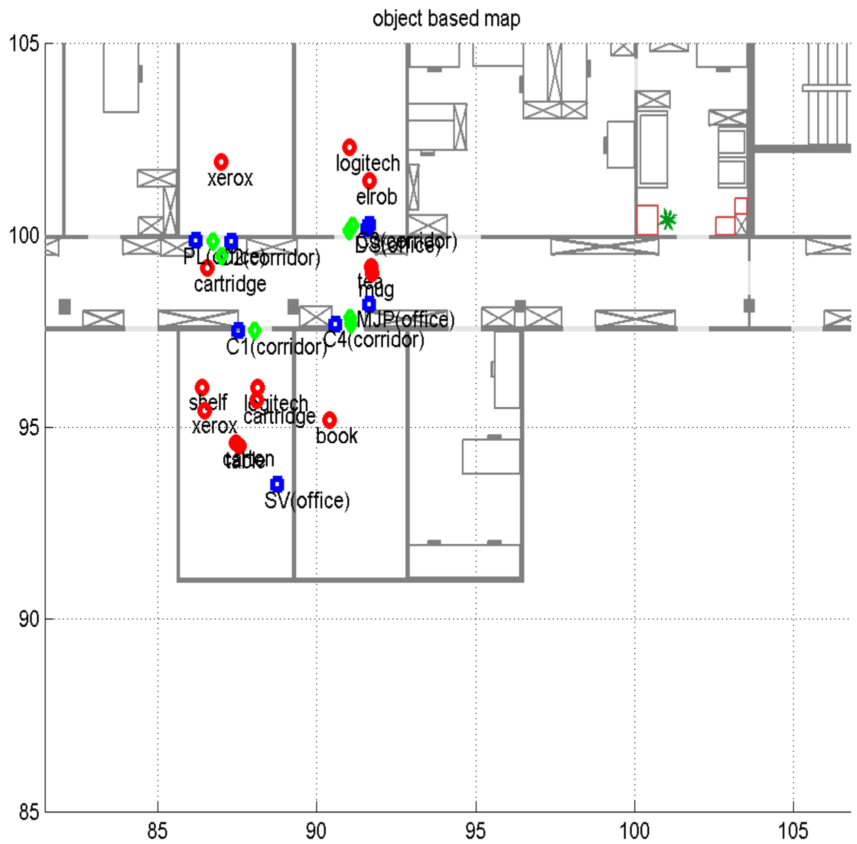
Scene 20 – in new place





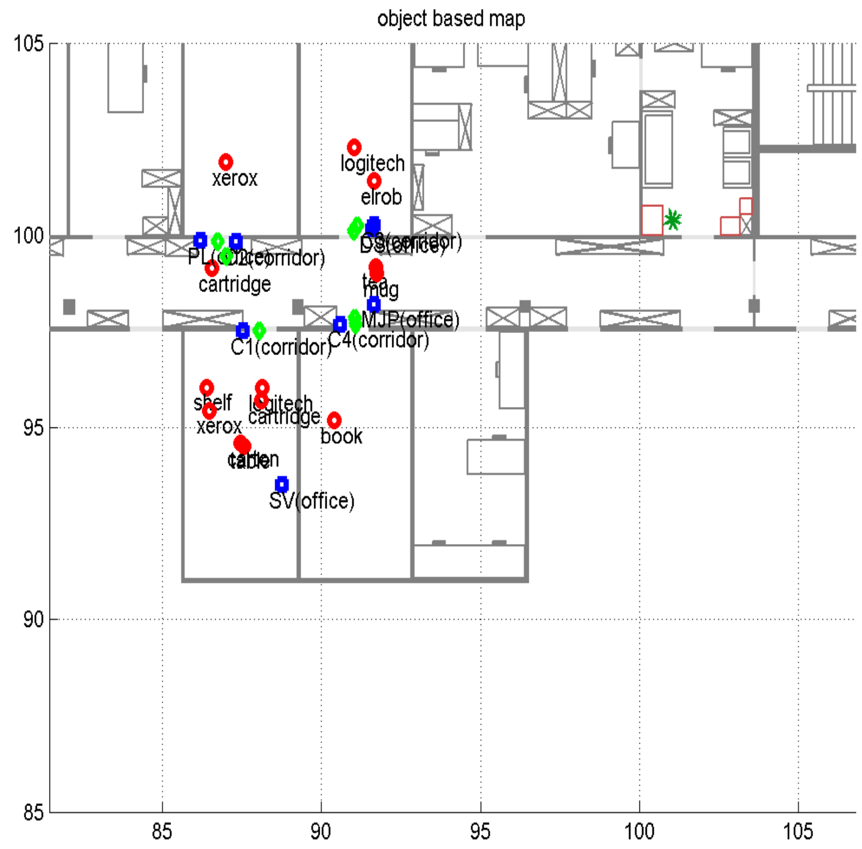
- object
- place reference
- door

Scene 22



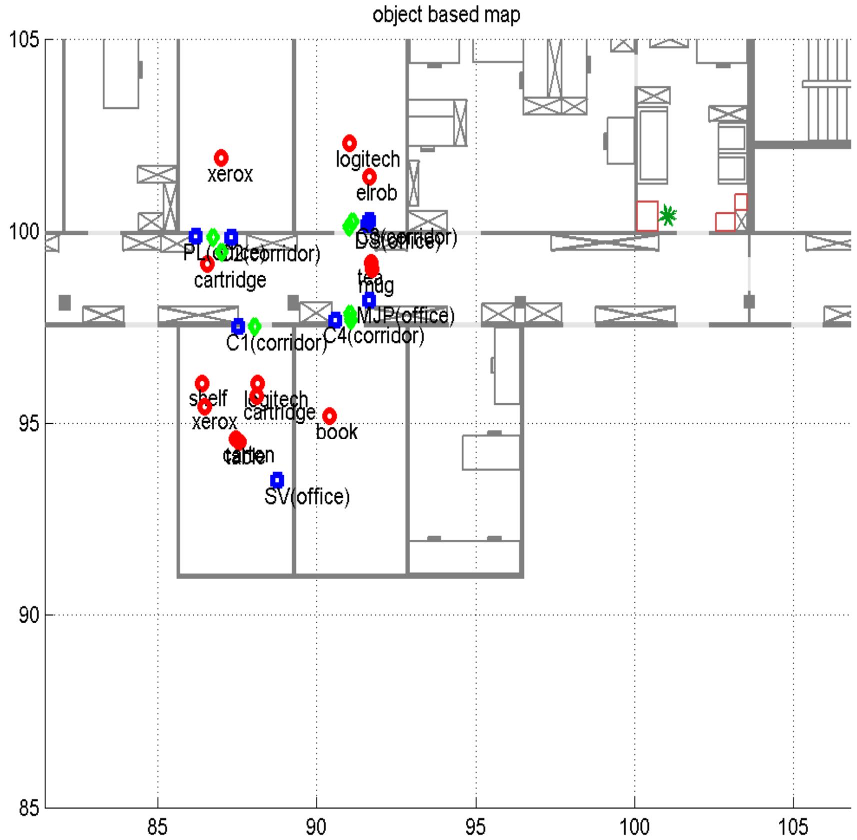
● object ● place reference ● door

Scene 23



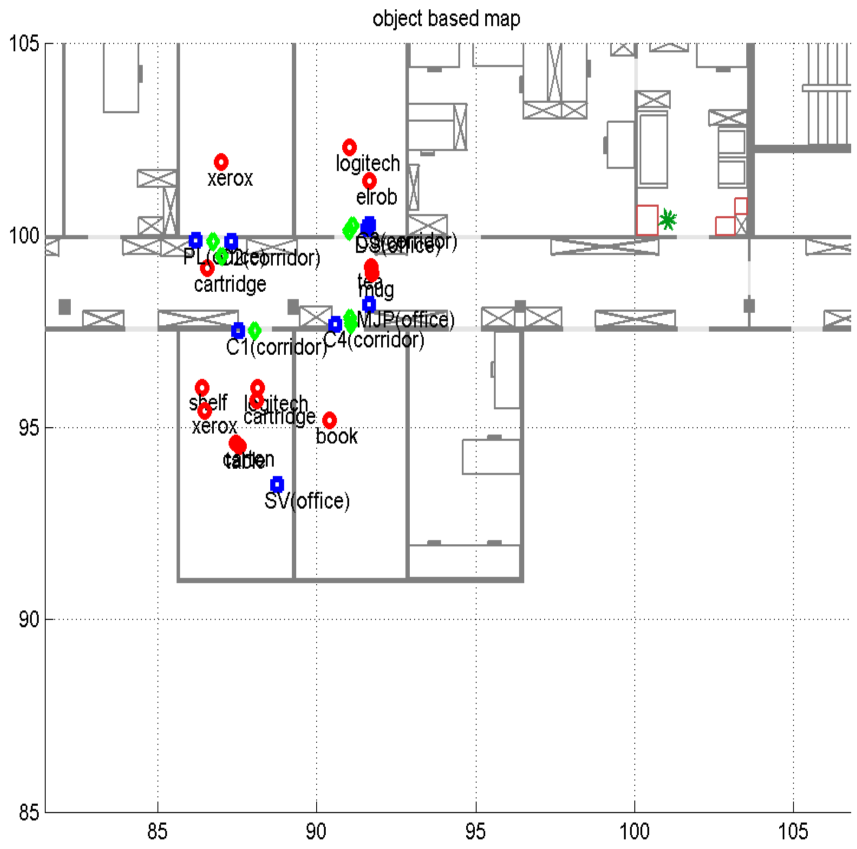
- object
- place reference
- door

Scene 24



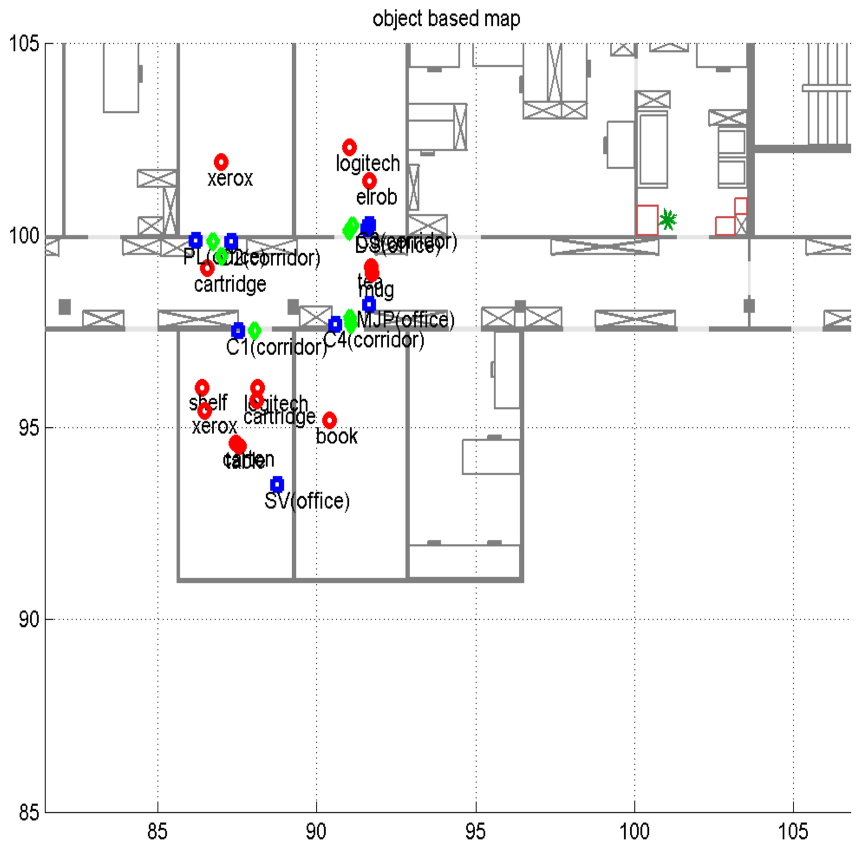
object place reference door

Scene 28



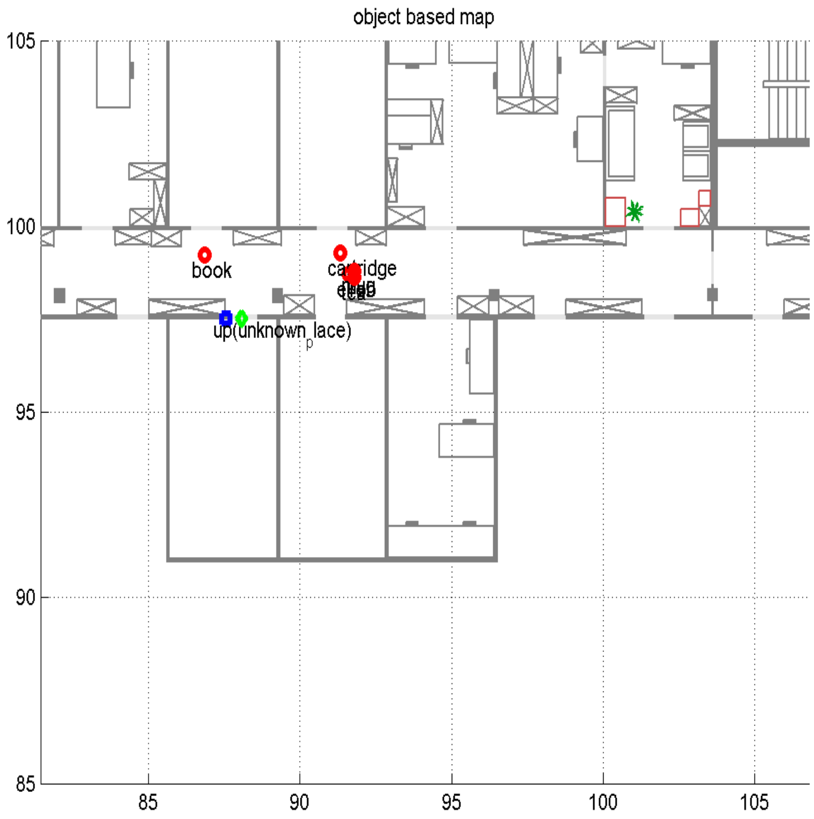
object place reference door

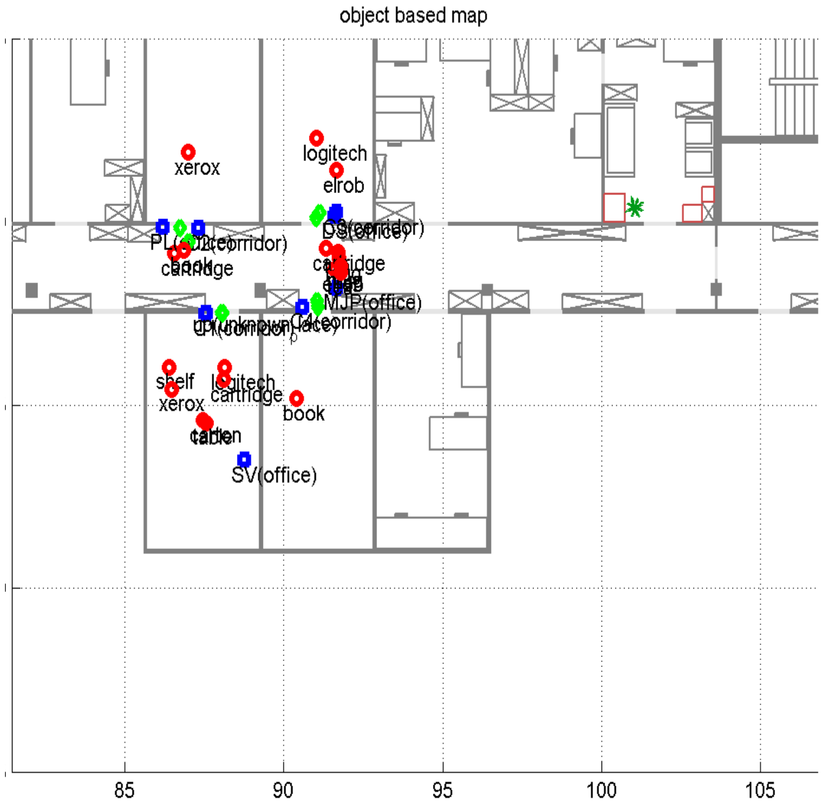
Scene 29



● object ● place reference ● door

Scene 30





Unknown place is classified as a corridor with a belief of 0.83. Recognition not possible with available data.

Current Work

- Integration / Mapping
  - Addressing consistency/ loop-closing issues , fusing of doorways, merging of places…
  - Testing
    - Larger data-sets ; robustness of approach.
    - Improving perception systems.
- Representation
  - What is a place ?
  - What kinds of semantic groups (ontologies) can be formed using the objects ?
  - What kinds of conditional independencies can be exploited to make the representation lighter ?
  - What other kinds of information could be incorporated ?

Acknowledgements

- Laurent Winkler, EPFL
- Davide Scaramuzza, ASL – EPFL
- Emanuele Frontoni, Università Politecnica delle Marche, Italy.
- COGNIRON

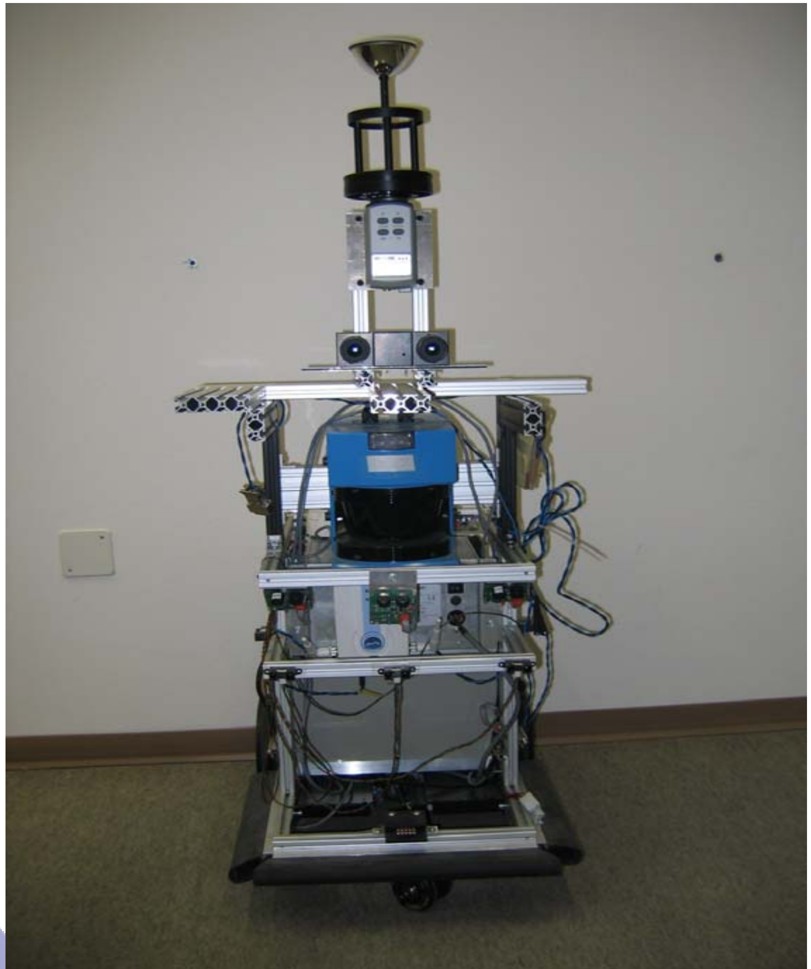
References

- Anguelov04
  - D. Anguelov, D. Koller, E. Parker, S. Thrun. **Detecting and Modeling Doors with Mobile Robots**. Proceedings of the International Conference on Robotics and Automation (ICRA), 2004.
- Arras03
  - Ph.D. Thesis. **Feature-Based Robot Navigation in Known and Unknown Environments**. Thèse N° 2765 (2003), Swiss Federal Institute of Technology, Lausanne.
- Choset01
  - H. Choset and K. Nagatani. **Topological simultaneous localization and mapping (SLAM): toward exact localization without explicit localization.** IEEE Transactions on Robotics and Automation, Volume: 17 Issue: 2 , Apr 2001.
- Galindo05
  - C. Galindo, A. Saffiotti, S. Coradeschi, P. Buschka, J.A. Fernández-Madrigal and J. González. **Multi-Hierarchical Semantic Maps for Mobile Robotics**. Proc. of the IEEE/RSJ Intl. Conf. on Intelligent Robots and Systems (IROS), pp. 3492-3497. Edmonton, CA, 2005.
- Kuipers83
  - B. J. Kuipers. **The cognitive map: Could it have been any other way?** Spatial Orientation: Theory, Research, and Application. New York: Plenum Press, 1983, pages 345-359.
- Kuipers00
  - B. Kuipers. **The Spatial Semantic Hierarchy**. Artificial Intelligence, 119: 191-233, May 2000.
- Lowe04
  - D. G. Lowe, **Distinctive image features from scale-invariant key-points**, International Journal of Computer Vision, vol. 60, no. 2, 2004.

References

- Martinelli04
  - A. Martinelli, A. Svensson, N. Tomatis and R. Siegwart. **SLAM Based on Quantities Invariant of the Robot's Configuration**. IFAC Symposyum on Inteligent Autonomous Vehicles, 2004.
- Stachniss05
  - C. Stachniss, Ó. Martínez-Mozos, A. Rottmann, and W. Burgard, **Semantic Labeling of Places**. In Proc. of the Int. Symposium of Robotics Research (ISRR).San Francisco, CA, USA, 2005.
- Tapus05
  - Ph.D. Thesis. **Topological SLAM : simultaneous localization and mapping with fingerprints of places** . Thèse N° 3357 (2005), Swiss Federal Institute of Technology, Lausanne.
- Thrun98
  - S. Thrun. **Learning metric-topological maps for indoor mobile robot navigation**. Artificial Intelligence, Volume 99, Issue 1, February 1998, Pages 21-71.
- Tomatis03
  - N. Tomatis, I. Nourbakhsh and R. Siegwart. **Hybrid Simultaneous Localization and Map Building: A Natural Integration of Topological and Metric.** Robotics and Autonomous Systems, 44, 3-14., July 2003.
- Torralba03
  - A. Torralba, K. P. Murphy, W. T. Freeman and M. A. Rubin., **Context-based vision system for place and object recognition**, Proceedings of the IEEE International Conference on Computer Vision, ICCV 2003, vol.1, p.273. Nice, France.
- Yeap01
  - W.K. Yeap and M.E. Jefferies, **On early cognitive mapping**. Spatial Cognition and Computation 2(2) 85-116, 2001.

Robot platform



Total run length < 20m

Usage of stereo / laser and odometry.

Stereo Error Model

- Based on work by Il-Kyun Jung and Simon Lacroix on SLAM using stereo vision.
- Equation:

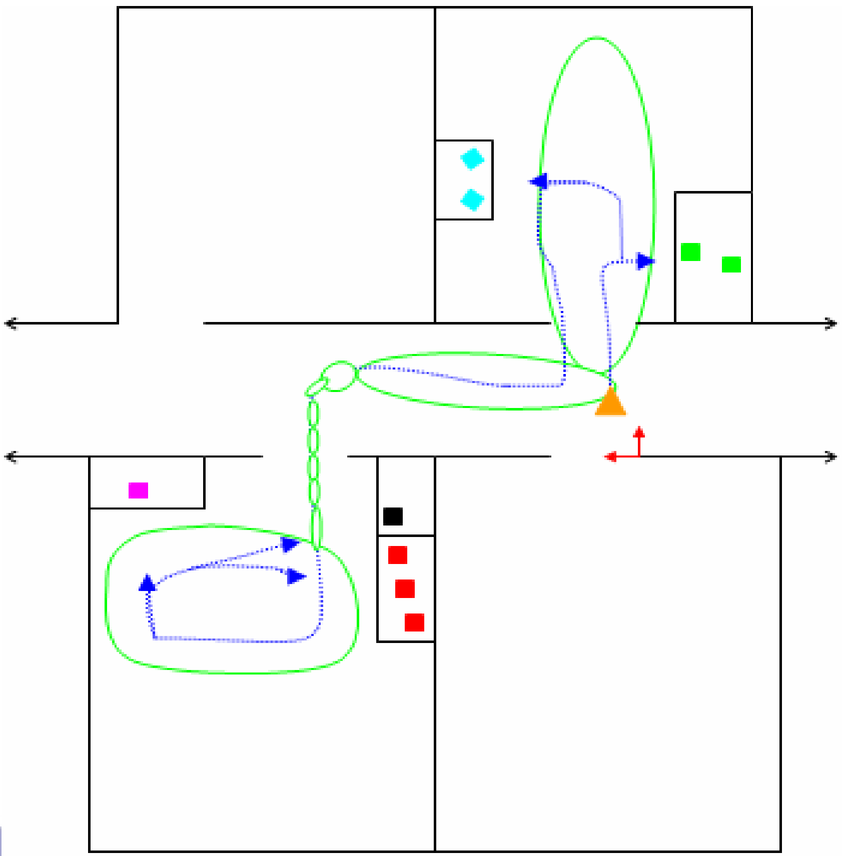
$$R_m = \begin{bmatrix} 1 & \beta_u & \gamma_u \\ \beta_u & \beta_u^2 & \beta_u \gamma_u \\ \gamma_u & \beta_u \gamma_u & \gamma_u^2 \end{bmatrix} \cdot \left(\frac{\sigma_d}{b.\alpha} . z^2 \right)^2$$

β & γ are obtained from the calibration process.

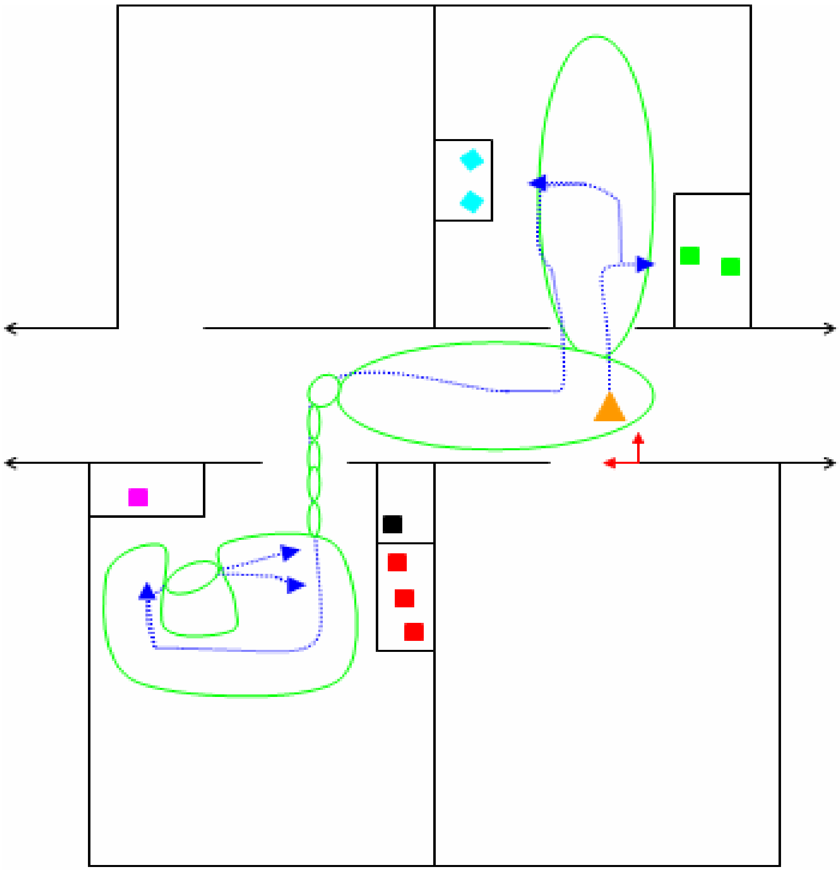
σ_δ is the standard deviation of the disparities

SIFT based place formation

Two way criterion



One way criterion



Notes

- Object coordinate selection
  - Using the affine transformation provided by the OR module – identify bounding box of object.
  - From the visible (disparity) area, estimate mean non-zero disparity.
  - Using the object center (from bounding box) and mean disparity, reproject object center in 3D camera space.
- Supervised learning / mapping process – the algorithms are capable of being applied in autonomous operation scenarios. The supervision allows for the incorporation of meaningful semantics.

More notes

- Why relative map?
  - To capture the relative spatial information between objects.
  - To incorporate invariant quantities to drift due to robot translation/rotation.