Reliable confirmation of an object identity by a mobile robot: A mixed appearance/localization-driven motion approach

Rafael Murrieta

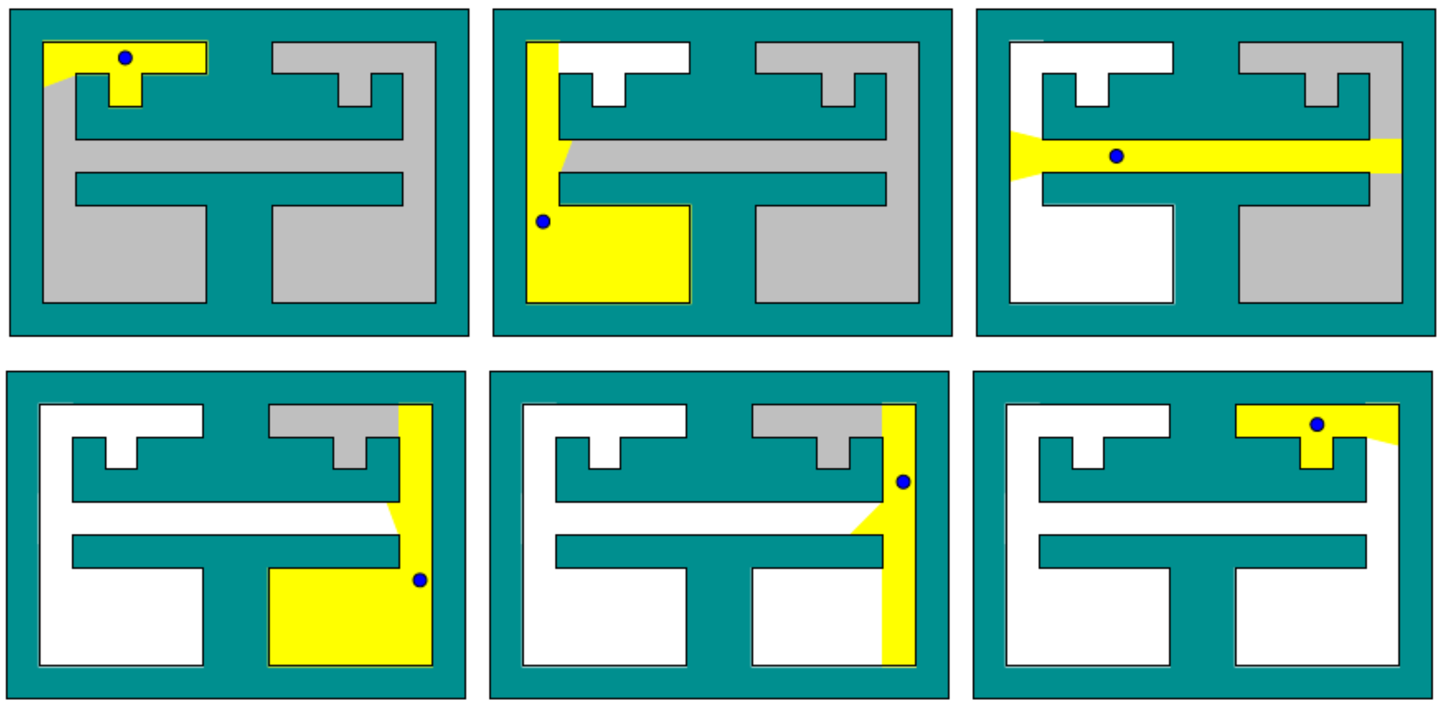
This is a joint work with: Israel Becerra, Luis M. Valentín and Jean-Claude Latombe

Outline

- Related work
- Problem definition
- Observation model
- Confirmation process
- Motion model
- Computation of motion strategy
- Simulation experiments
- Real-world experiments
- Conclusions

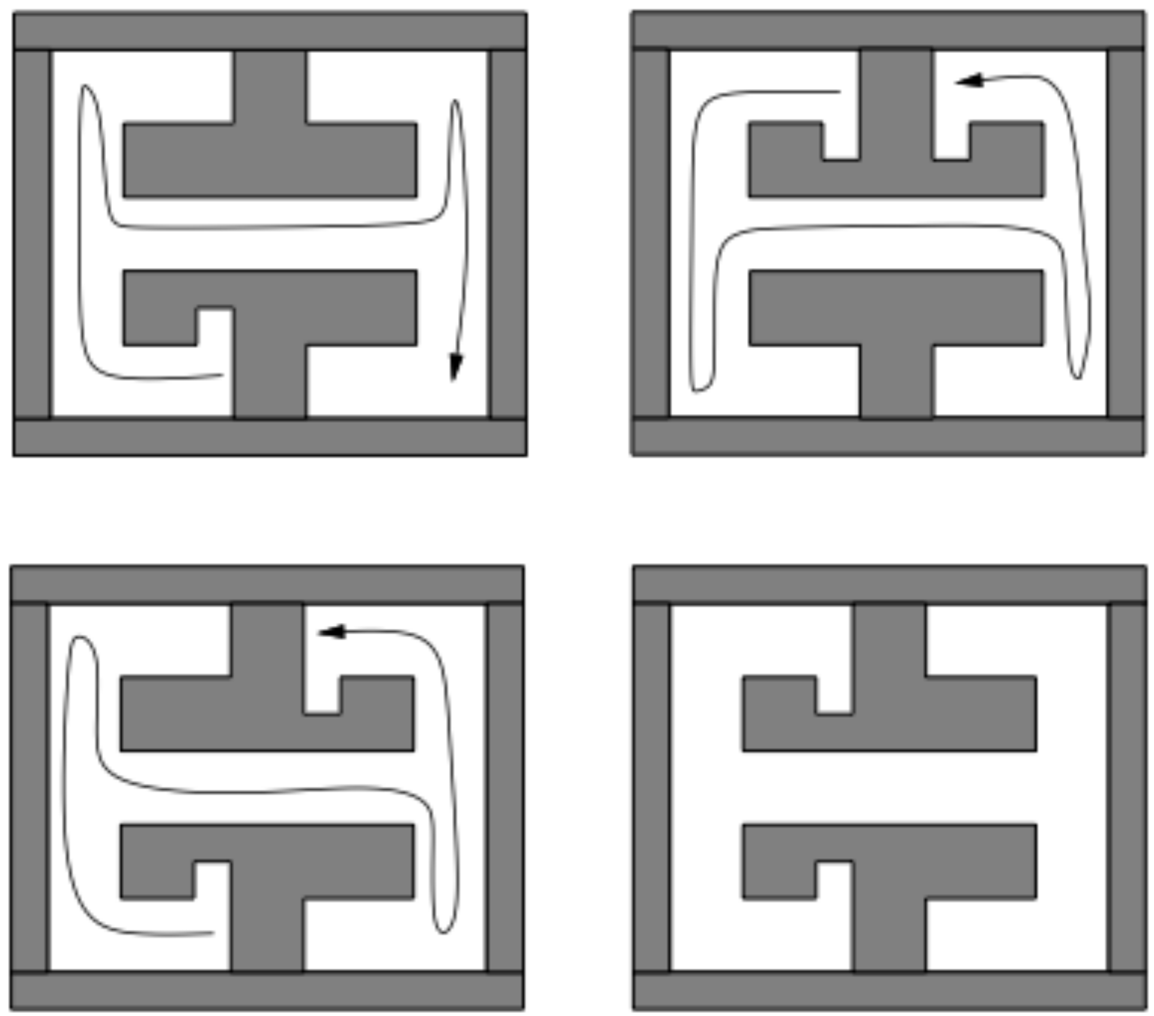
Related work

- Target ***finding*** in an environment with obstacles
    - V. Isler et. al., “Randomized pursuit-evasion in a polygonal enviroment”, IEEE Transactions on Robotics, vol. 5, no. 21, pp. 864-875, 2005.
    - R. Vidal et. al., “Probabilistic pursuit-evasion games: Theory, implementation, and experimental evaluation”, IEEE Transactions on Robotics and Automation, vol. 18, no. 5, pp.662-669, 2002.
    - L. Guibas, J.-C. Latombe, S. LaValle, D. Lin and R. Motwani, “Visibility-based pursuit-evasion in a polygonal environment”, International Journal of Computational Geometry and Applications, vol. 9, no. 4/5, pp. 471-494, 1999.



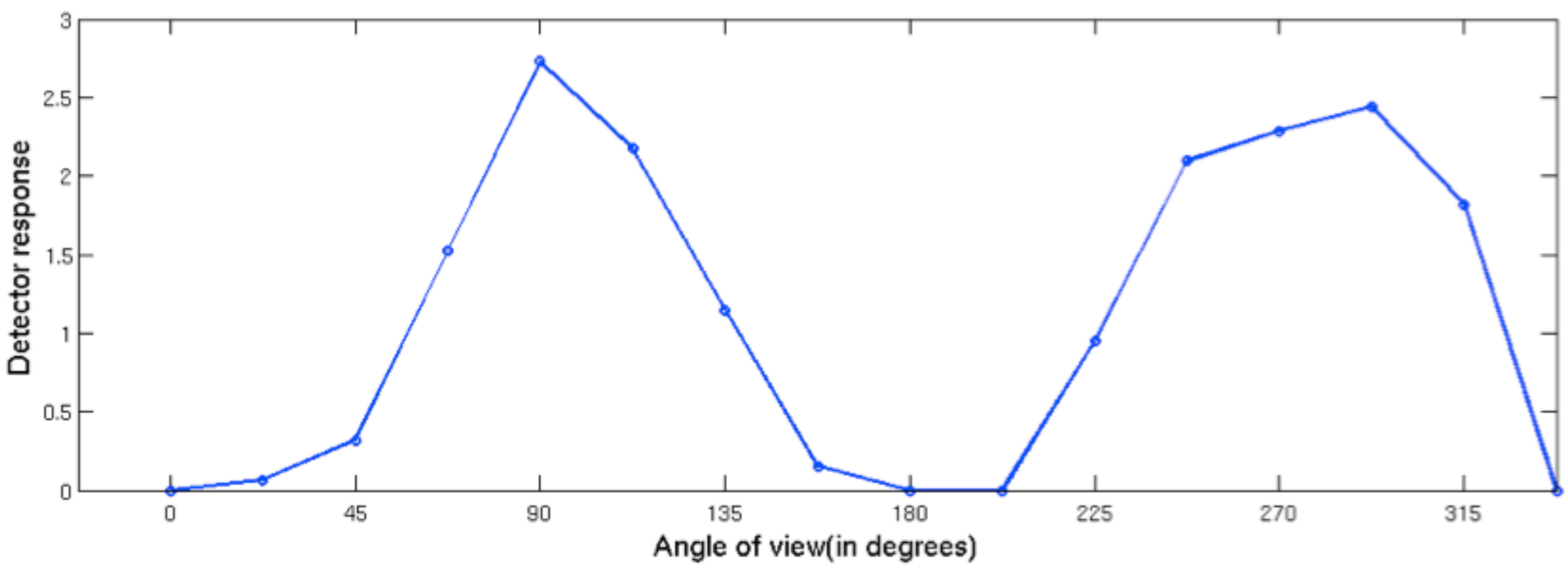
Related work

L. Guibas, J. Latombe, S. Lavalle, D. Lin and R. Motwani, "Visibility-based pursuit-evasion in a polygonal environment", *International Journal of Computational Geometry and Applications*, pp. 17-30, 1997.



Related work

- D. Meger, A. Gupta and J. Little, "Viewpoint Detection Models for Sequential Embodied Object Category Recognition", *ICRA*, 2010.



Related work

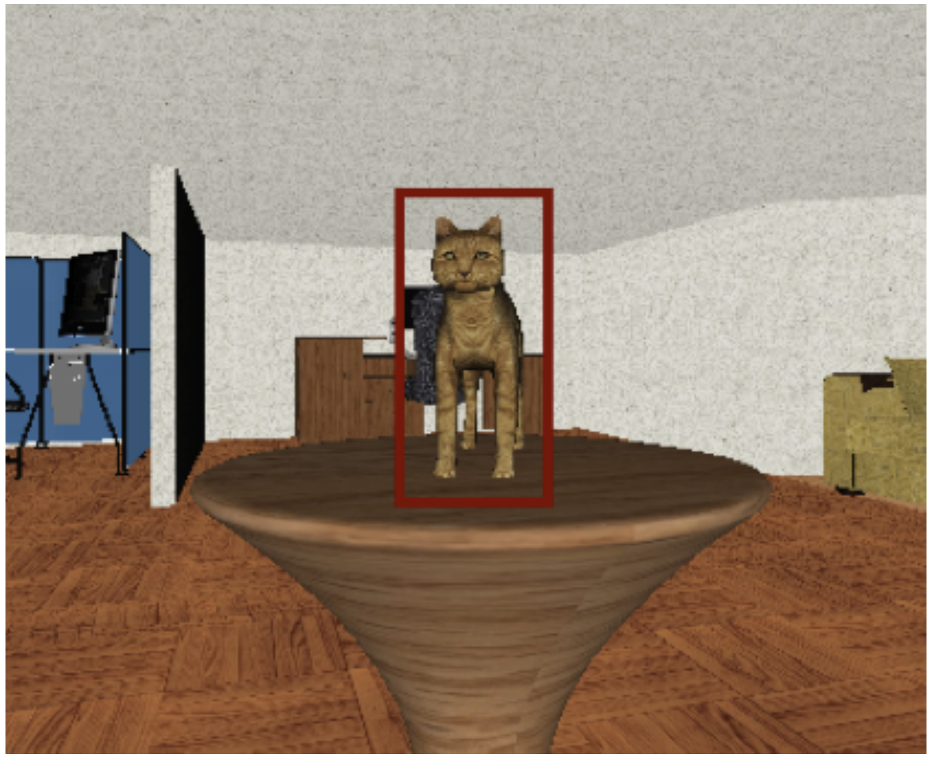
- [1] J. Espinoza, A. Sarmiento, R. Murrieta-Cid and S. Hutchinson, “A Motion Planning Strategy for Finding an Object with a Mobile Manipulator in 3-D Environments”, *Advanced Robotics*, 2011.

Problem Definition

- We investigate an object detection problem using a mobile robot equipped with a vision sensor.
- The robot is instructed to find a certain object **T** (the target) in its environment.
- At some point of the search process, the robot believes that it has encountered a candidate **C** for **T**, but it is not sure yet that **C** is **T** (module in [1]).
- Taking advantage of its mobility, the robot tries to achieve adequate viewpoints to confirm (or infirm) that **C** is actually **T** .

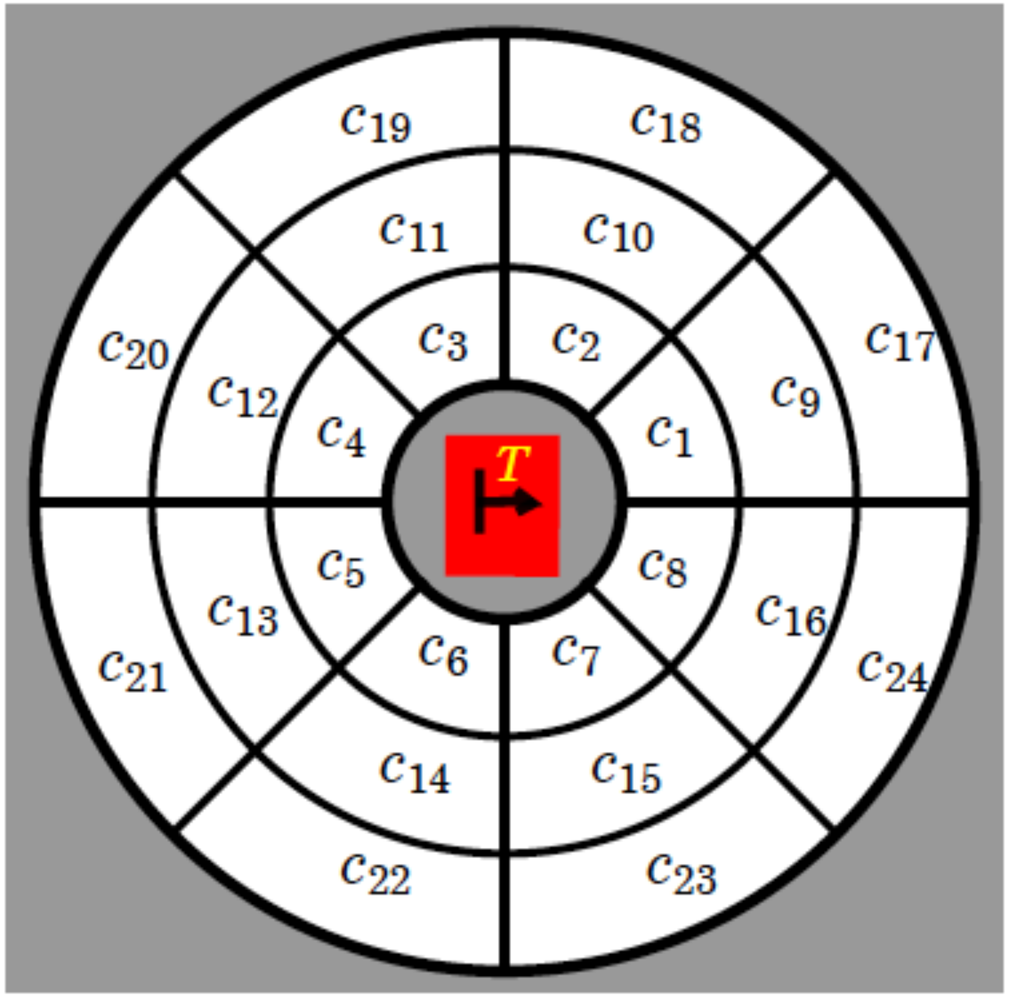
Observation model

- The robot is equipped with a software module DT (detector), capable of identifying **T**.



- DT returns a discrete detection score $y \in \{o_1, o_2, .., o_n\}$, where $o_1 < o_2 < ... < o_n$, measuring how well the image matches the appearance of **T**, hence the confidence of the identification.

Observation model



- The observation model of T is then created in the form of a probability distribution $P(o_j|c_i)$.

Confirmation process

- The location xt of the robot will be modeled at the cell resolution by a probability distribution over the m cells of X

$$P(x_t|I_t) = \frac{\sum_{x_{t-1}} P(x_{t-1}|I_{t-1})P(x_t|x_{t-1},u_{t-1})P(y_t|x_t)}{\sum_{x_t}\sum_{x_{t-1}} P(x_{t-1}|I_{t-1})P(x_t|x_{t-1},u_{t-1})P(y_t|x_t)}$$

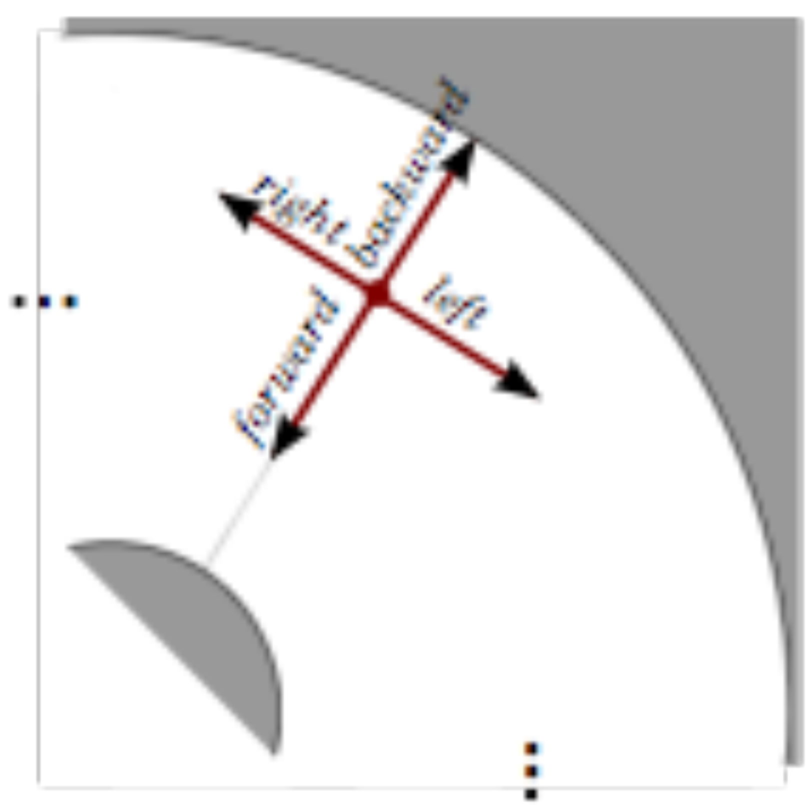
Confirmation process

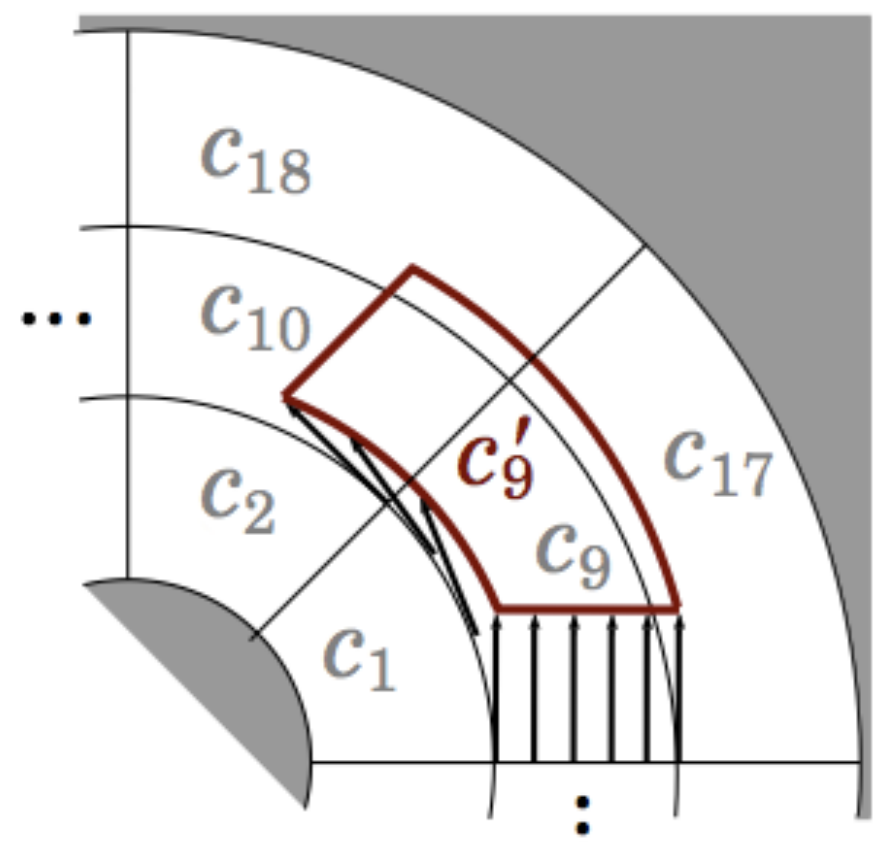
- The target is declared as detected if the detector returns a confidence score greater than ô at time t +1 and If the robot reaches at time t a position where the condition $P(y_{t+1} \geq \hat{o} | I_t, u_t) > \lambda$ is satisfied.

- *This gives us a twofold goal that mixes robot localisation relatively to the candidate object and target identification using its appearance.*

Motion model

- The motion model is given by the probability distribution $P(x_t|x_{t-1}, u_{t-1})$.
- We have 4 motion commands .





Computation of motion strategy $\pi(t, I_t)$

- We use SDP to calculate the motion policy

$$J_{N-1}(I_{N-1}) = \max_{u_{N-1} \in U_{N-1}} \Big[\tilde{g}(I_{N-1}, u_{N-1}) + \mathop{E}_{x_{N-1}} \Big\{ \mathop{E}_{x_N} \{ g_F(x_N) | x_{N-1}, u_{N-1} \} | I_{N-1}, u_{N-1} \Big\} \Big]$$

$$\pi(N-1, I_{N-1}) = \arg \max_{u_{N-1} \in U_{N-1}} \Big[\tilde{g}(I_{N-1}, u_{N-1}) + \mathop{E}_{x_{N-1}} \Big\{ \mathop{E}_{x_N} \{ g_F(x_N) | x_{N-1}, u_{N-1} \} | I_{N-1}, u_{N-1} \Big\} \Big]$$

and for $t < N - 1$

$$J_t(I_t) = \max_{u_t \in U_t} \Big[\tilde{g}(I_t, u_t) + \mathop{E}_{y_{t+1}} \{ J_{t+1}(I_t, y_{t+1}, u_t) | I_t, u_t \} \Big]$$

$$\pi(t, I_t) = \arg \max_{u_t \in U_t} \Big[\tilde{g}(I_t, u_t) + \mathop{E}_{y_{t+1}} \{ J_{t+1}(I_t, y_{t+1}, u_t) | I_t, u_t \} \Big]$$

Computation of motion strategy

- We use SDP to calculate the motion policy $\pi(t, I_t)$.
- Since we want the robot to achieve a position where holds, we set the gain function $\tilde{g}(I_t, u_t)$ to:

$$P(y_{t+1} \geq \hat{o} | I_t, u_t) > \lambda$$

$$P(y_{t+1} \geq \hat{o} | I_t, u_t) = \sum_{x_{t+1}} P(y_{t+1} \geq \hat{o} | x_{t+1}) \sum_{x_t} P(x_{t+1} | x_t, u_t) P(x_t | I_t)$$

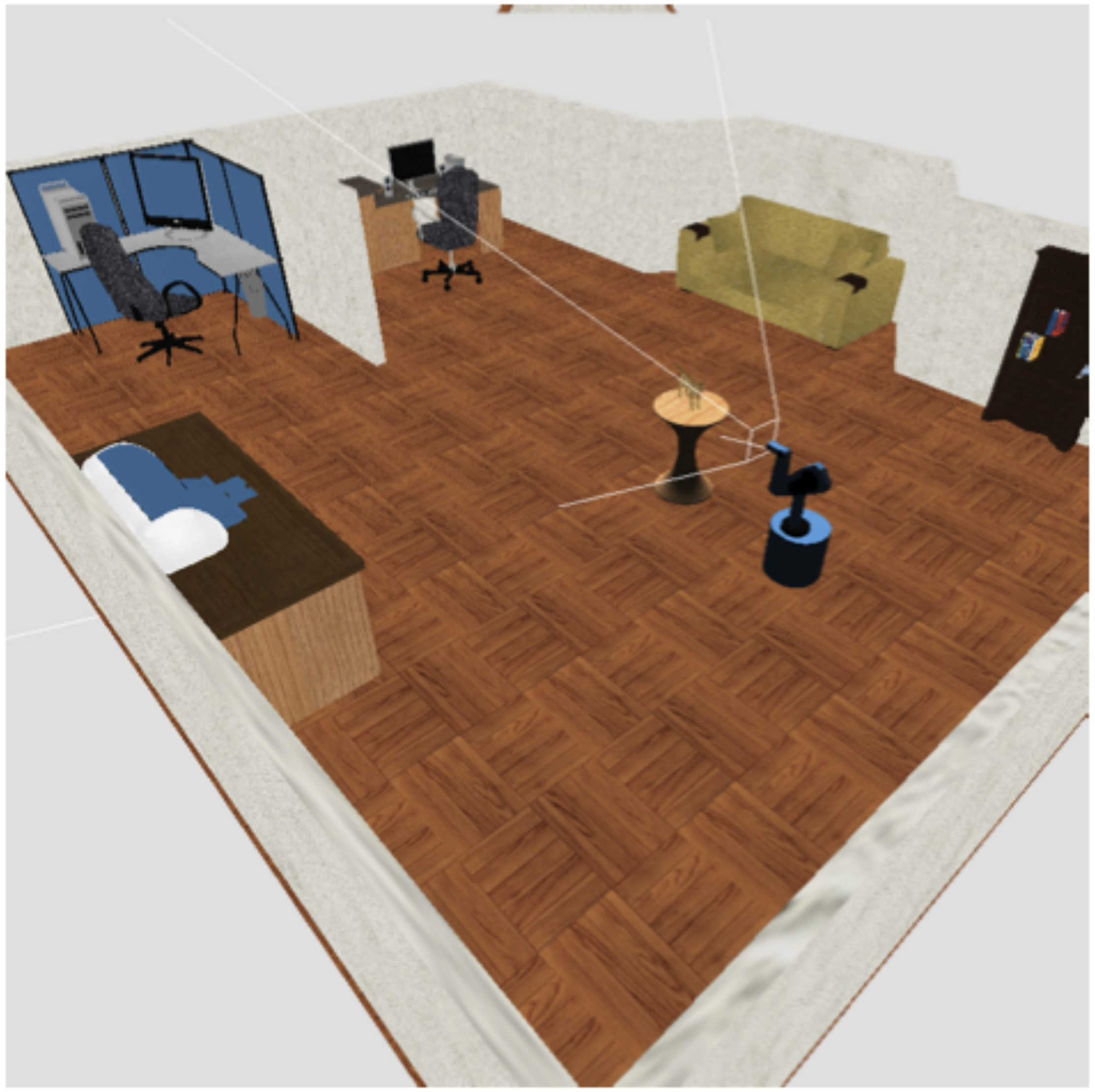
Computation of motion strategy

- If the robot reaches at time t a position where the condition $P(y_{t+1} \geq \hat{o} | I_t, u_t) > \lambda$ is satisfied and if the detector returns a confidence score greater than ô at time t + 1, then the confirmation process ends with success; otherwise, a new N horizon strategy is again computed.

- This iterative process ends whenever the goal is achieved or when $p \times N \geq Ne$, in which case the robot considers that the object **C** is not the target **T**

Simulation experiments

- Virtual environment



Simulation experiments

- We use a 24-cell decomposition.

- For each target **T**, the detector DT uses a deformable part model algorithm [2] trained on a set of images taken from a single cell cg of the decomposition.

  [2] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models", Trans. on Pattern Analysis and Machine Intelligence, 2010.

- 6 score values as observation.

Exp #1

- Target

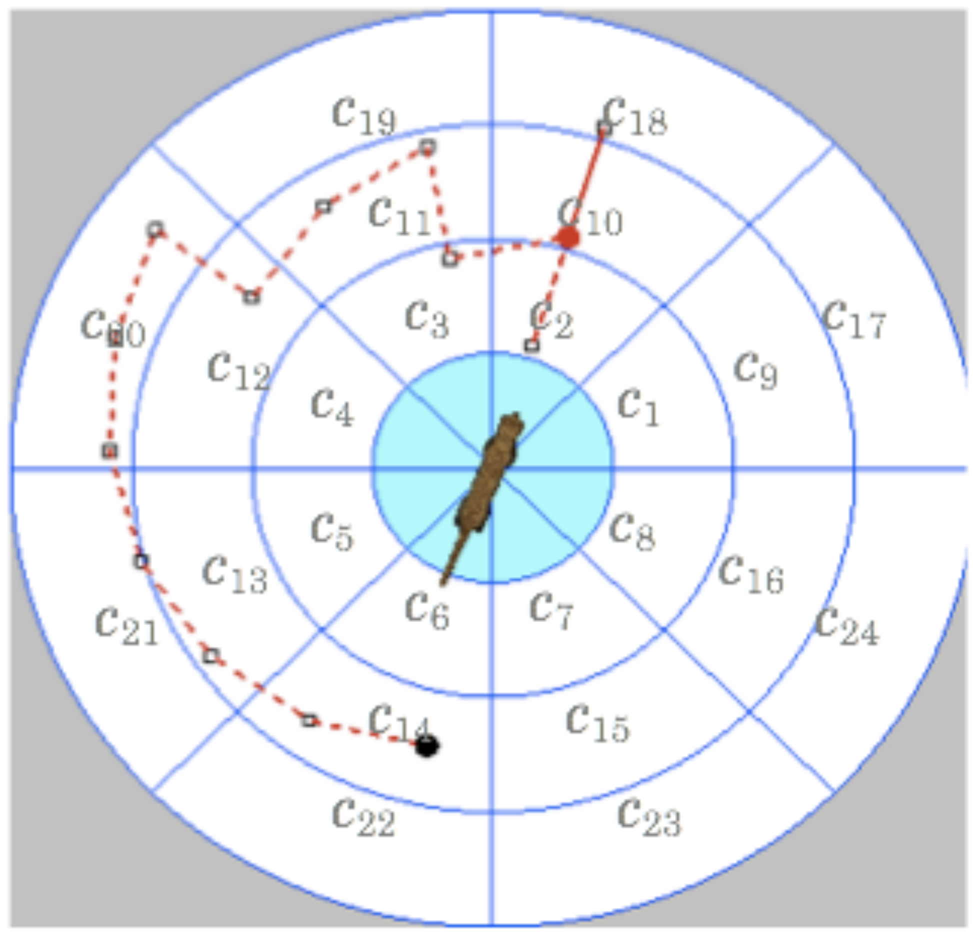
Exp #1

- Cat´s observation model

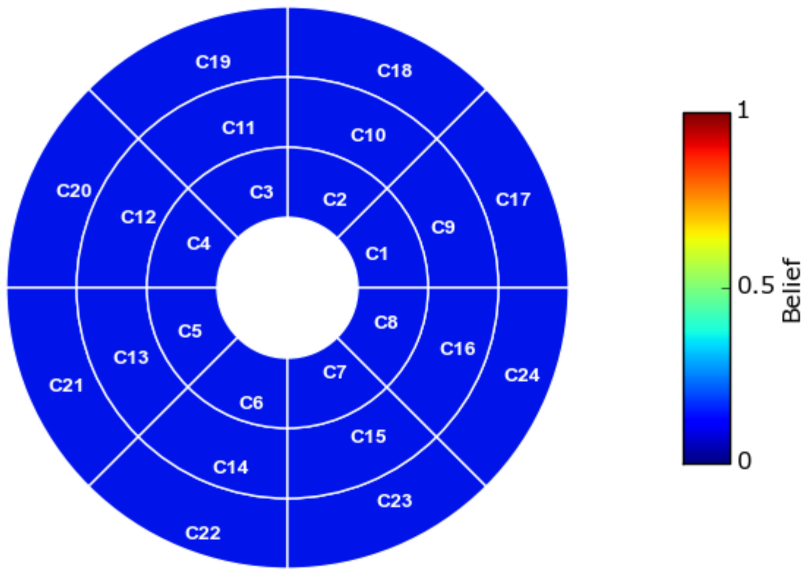
	o_1	o_2	o_3	o_4	o_5	o_6
c_1	0.236	0.324	0.220	0.140	0.068	0.012
c_2	0.000	0.000	0.000	0.132	0.644	0.224
c_3	0.196	0.376	0.224	0.160	0.044	0.000
c_4	0.916	0.084	0.000	0.000	0.000	0.000
c_5	0.312	0.560	0.128	0.000	0.000	0.000
c_6	0.000	0.876	0.124	0.000	0.000	0.000
c_7	0.144	0.688	0.168	0.000	0.000	0.000
c_8	1.000	0.000	0.000	0.000	0.000	0.000
c_9	0.152	0.412	0.156	0.180	0.092	0.008
c_{10}	0.000	0.000	0.000	0.000	0.428	0.572
c_{11}	0.112	0.360	0.268	0.096	0.144	0.020
c_{12}	0.896	0.104	0.000	0.000	0.000	0.000
c_{13}	0.276	0.616	0.108	0.000	0.000	0.000
c_{14}	0.000	0.420	0.580	0.000	0.000	0.000
c_{15}	0.200	0.540	0.260	0.000	0.000	0.000
c_{16}	1.000	0.000	0.000	0.000	0.000	0.000
c_{17}	0.596	0.204	0.124	0.068	0.008	0.000
c_{18}	0.000	0.040	0.320	0.372	0.260	0.008
c_{19}	0.440	0.244	0.232	0.064	0.020	0.000
c_{20}	1.000	0.010	0.290	0.390	0.290	0.010
c_{21}	0.804	0.196	0.000	0.000	0.000	0.000
c_{22}	0.068	0.880	0.052	0.000	0.000	0.000
c_{23}	0.740	0.260	0.000	0.000	0.000	0.000
c_{24}	1.000	0.000	0.000	0.000	0.000	0.000

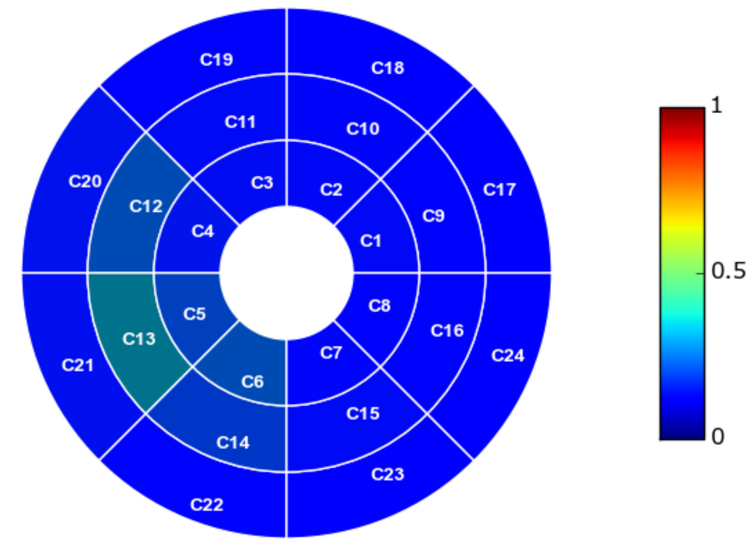
Simulation results-Exp #1

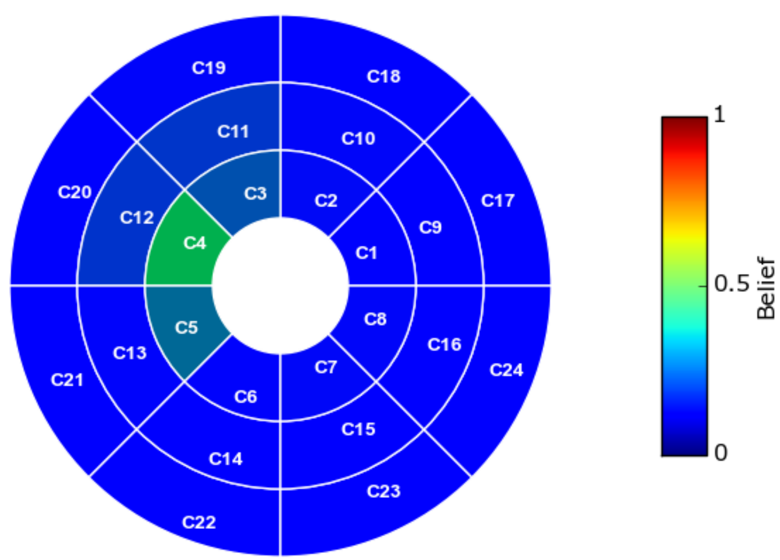
Path with N=3 and lambda=0.55

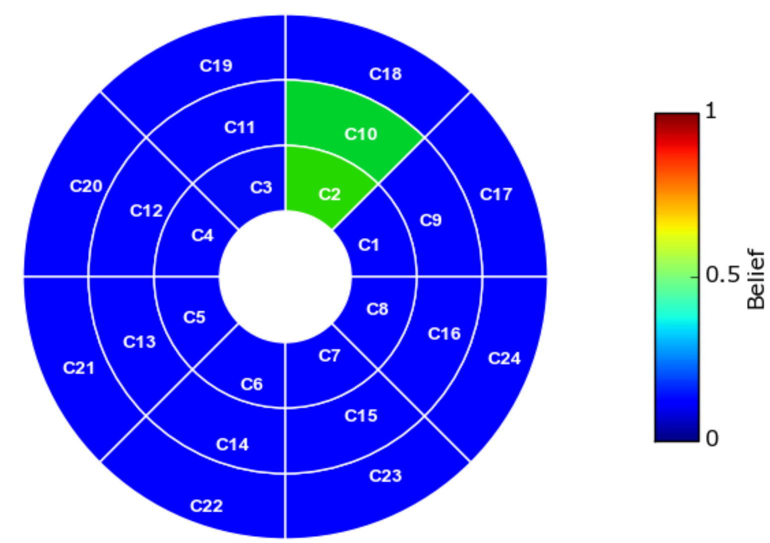


Evolution of the probability distribution P(Xt|It)



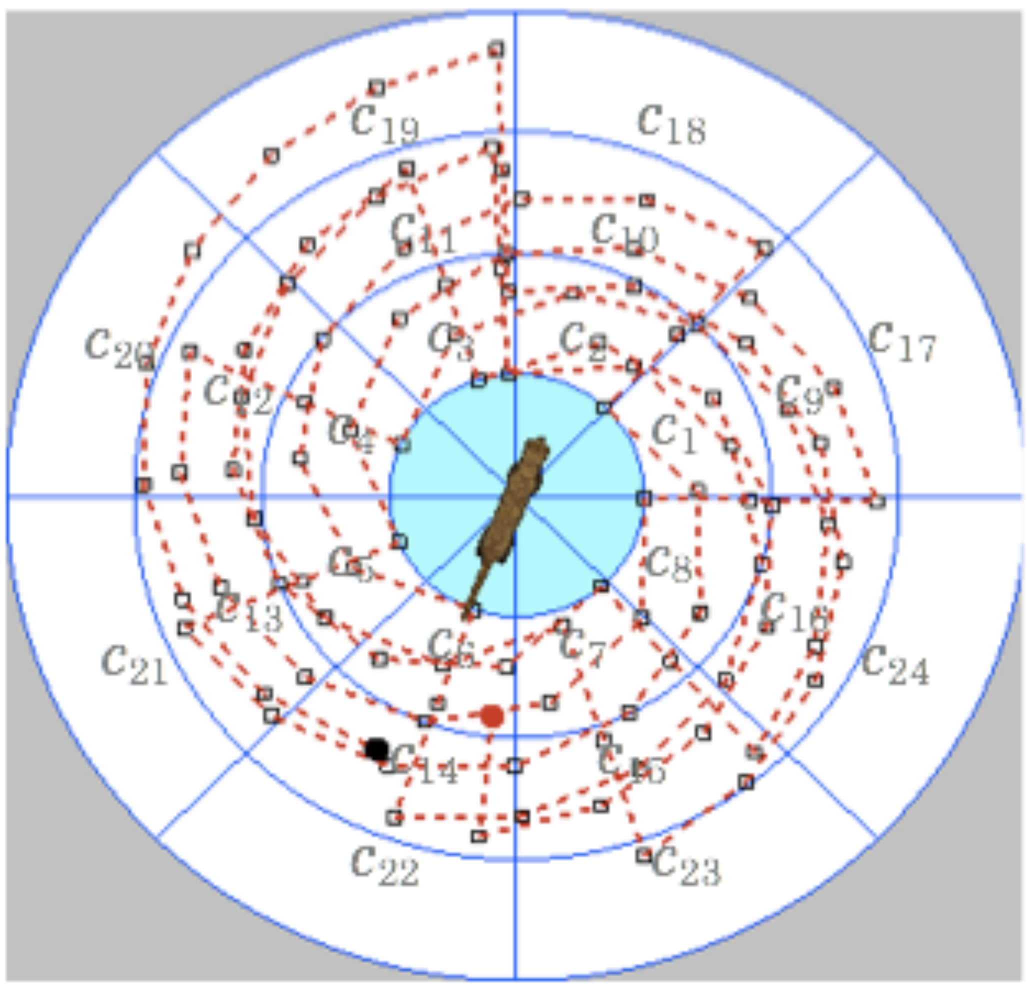


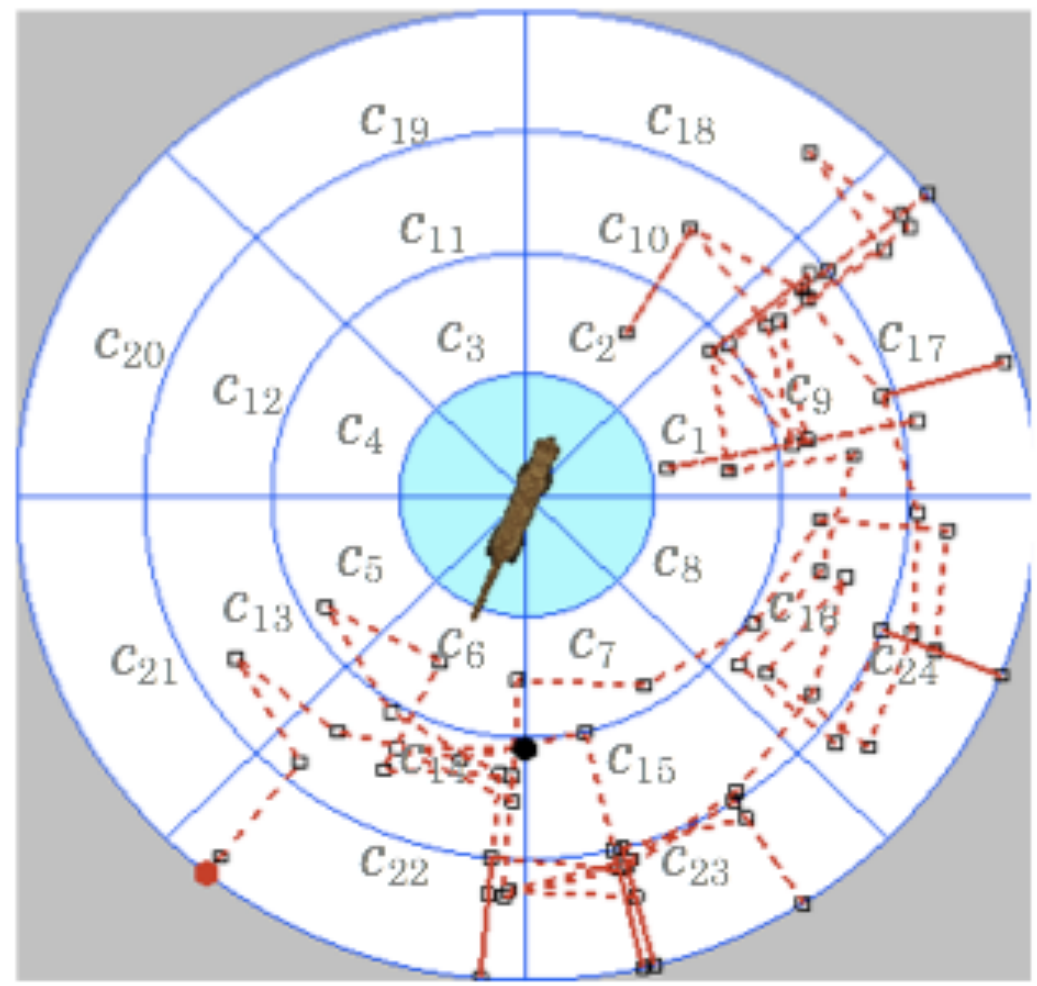




Simulation results-Exp #1

- Paths with lambda=0.55



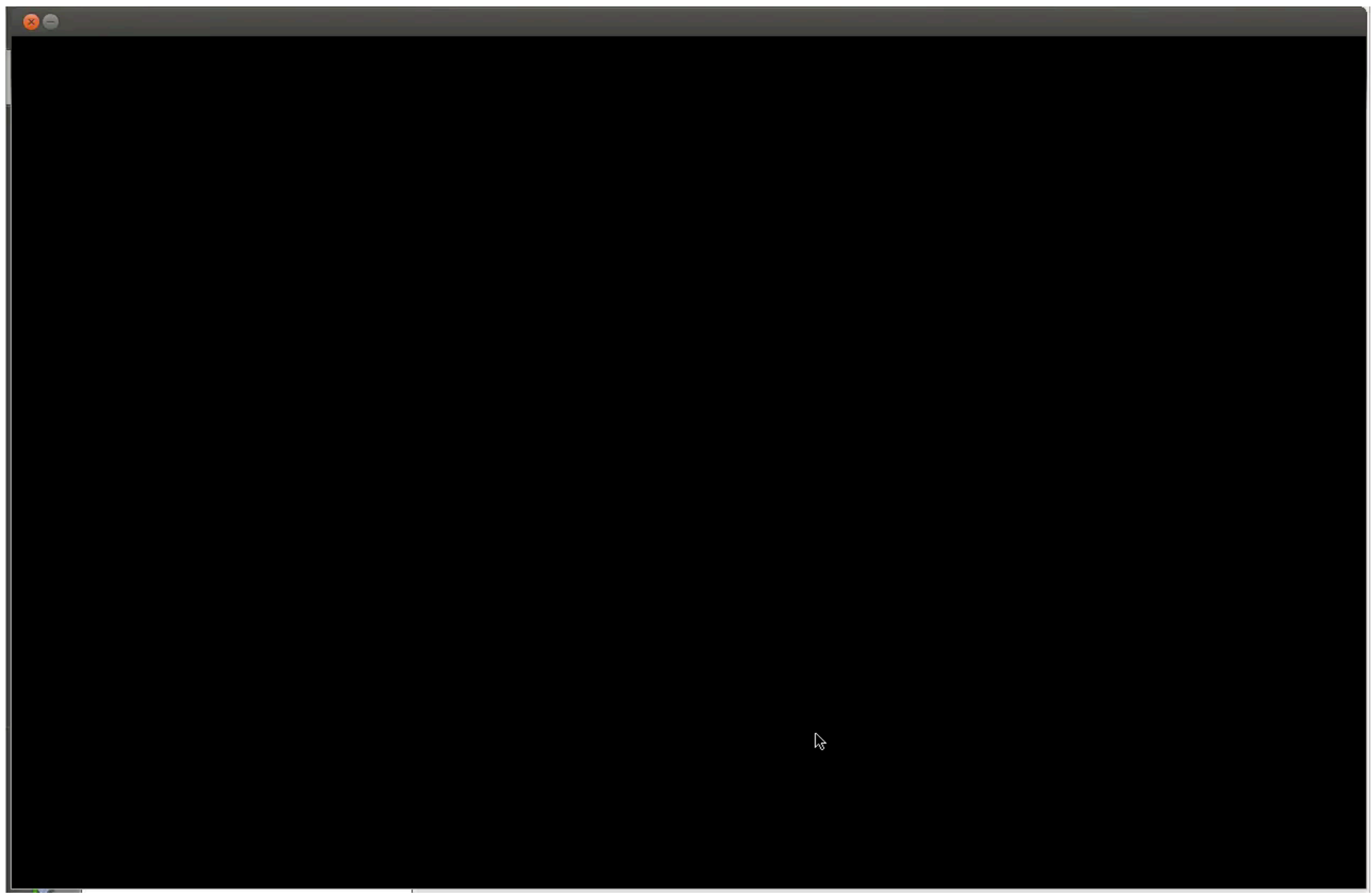


Exp #1

- Statistics

Planning horizon	λ	# of sensing locations	Path length	Planning time (ms)	% of confirmation
1	0.45	33.698	32.580	9.854	93
	0.50	42.994	41.837	12.693	88
	0.55	42.493	41.385	12.513	88
2	0.45	12.915	11.816	37.428	100
	0.50	13.150	12.068	38.004	100
	0.55	14.385	13.305	42.443	100
3	0.45	12.837	11.471	405.655	100
	0.50	13.120	11.715	415.278	100
	0.55	13.875	12.385	440.959	100
4	0.45	12.230	11.028	35485.734	100
	0.50	12.587	11.402	37040.285	100
	0.55	13.655	12.431	40236.170	100
Random	0.45	55.750	43.344	-	22
	0.50	62.962	46.266	-	26.5
	0.55	61.354	47.201	-	15.5

Exp #1



Exp #2

- Target

(a) True bottle (b) False bottle

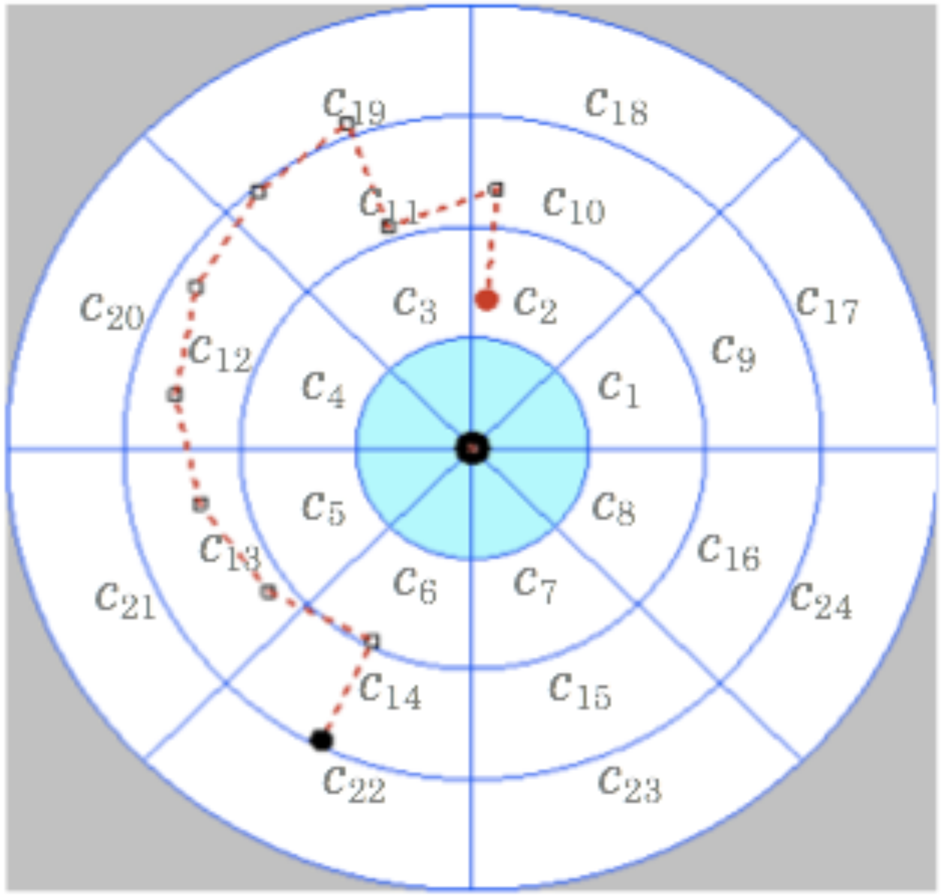
Exp #2

- Statistics

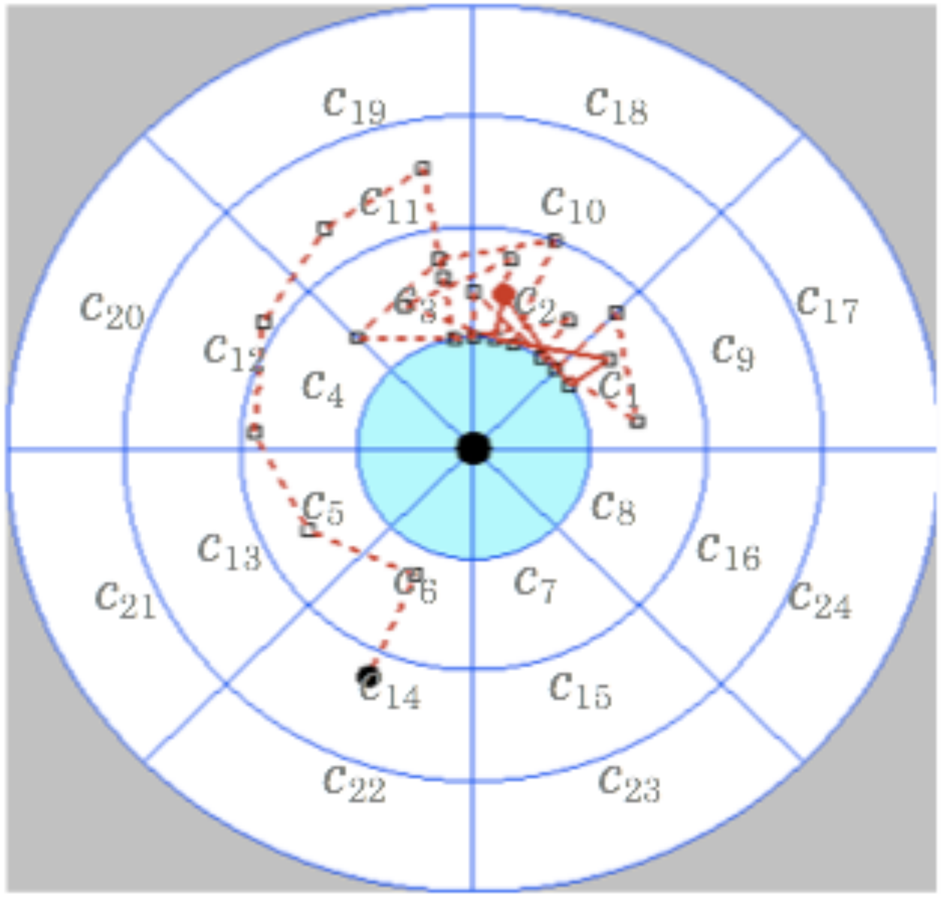
Scene object	λ	# of sensing locations	Path length	Planning time (ms)	% of confirmation
True Bottle	0.80	10.820	9.346	367.723	100
	0.85	10.825	9.122	361.993	100
	0.90	12.030	9.244	415.965	99.5
False Bottle	0.80	21.333	18.002	721.861	1.5
	0.85	17	14.561	621.074	0.5
	0.90	-	-	-	0

Exp #2

- Paths



(c) Path generated with $N = 3$ and $\lambda = 0.8$ (true bottle)



(d) Path generated with $N = 3$ and $\lambda = 0.8$ (false bottle)

Simulation results-Exp #3

- Target

Simulation results-Exp #3

- We compare N=2 and N=3, in 200 runs, with lambda=0.85, Ne=30 and DT trained at cell 10
- N=3 manage to confirm detection in 100% of the runs
- N=2 gets trapped in a local maximum at cell 5, only confirming detection in 1% of the runs

Simulation results-Exp #3

- Solution – we make gF a concave function with a maximum in cell cg (here, c10), so that it gives an incentive for the robot to reach cells in the neighborhood of cg

$$
\begin{aligned}
J_{N-1}(I_{N-1}) = \max_{u_{N-1}\in U_{N-1}} \Bigg[& \tilde{g}(I_{N-1}, u_{N-1}) \\
& + \underset{x_{N-1}}{E} \left\{ \underset{x_N}{E} \{ g_F(x_N) | x_{N-1}, u_{N-1} \} | I_{N-1}, u_{N-1} \right\} \Bigg]
\end{aligned}
$$

Imperfect knowledge of the table's center

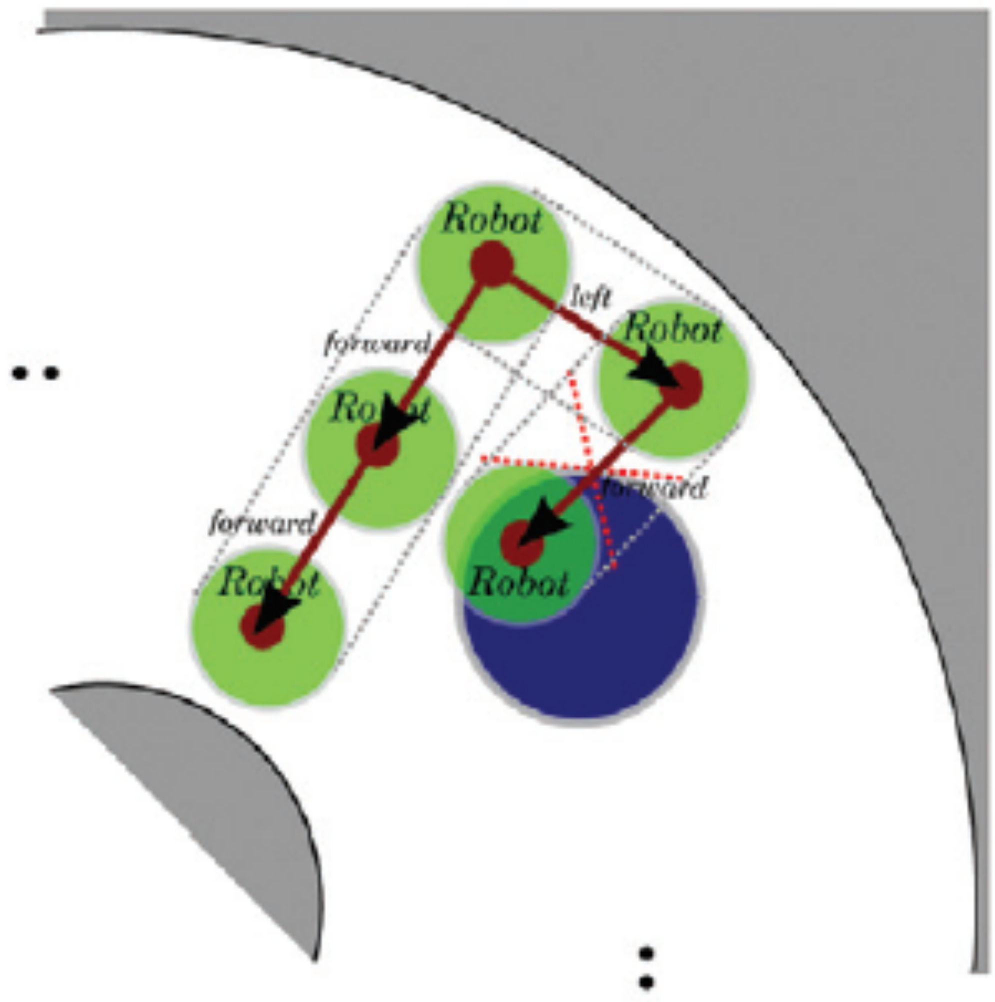
(a) Cat as target.

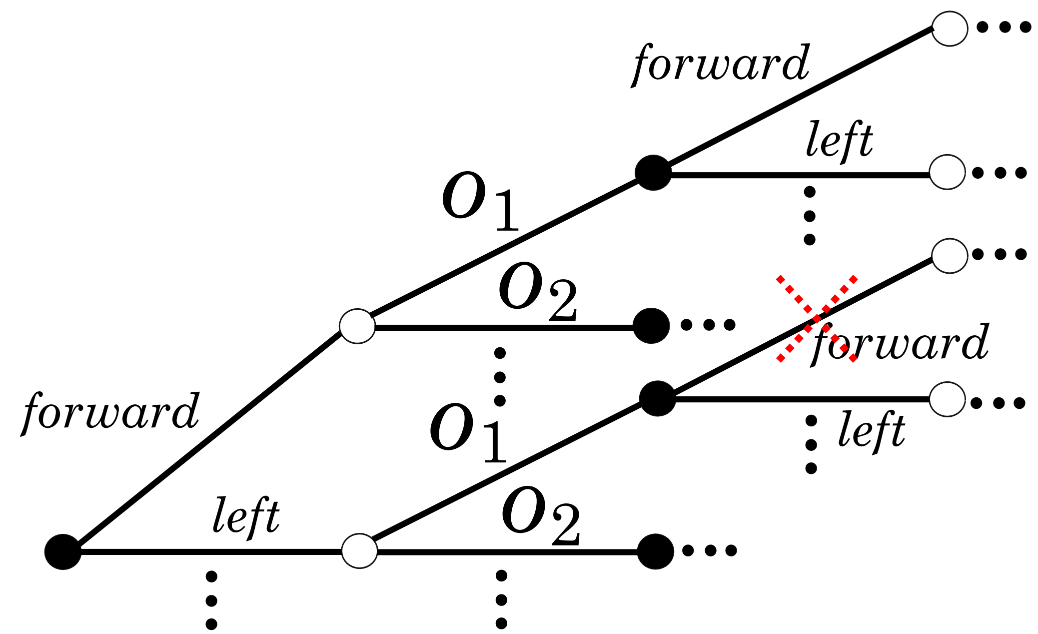
(b) Bottle as target.

Table 6. Statistics for Experiment #4 (error in robot alignment towards C).

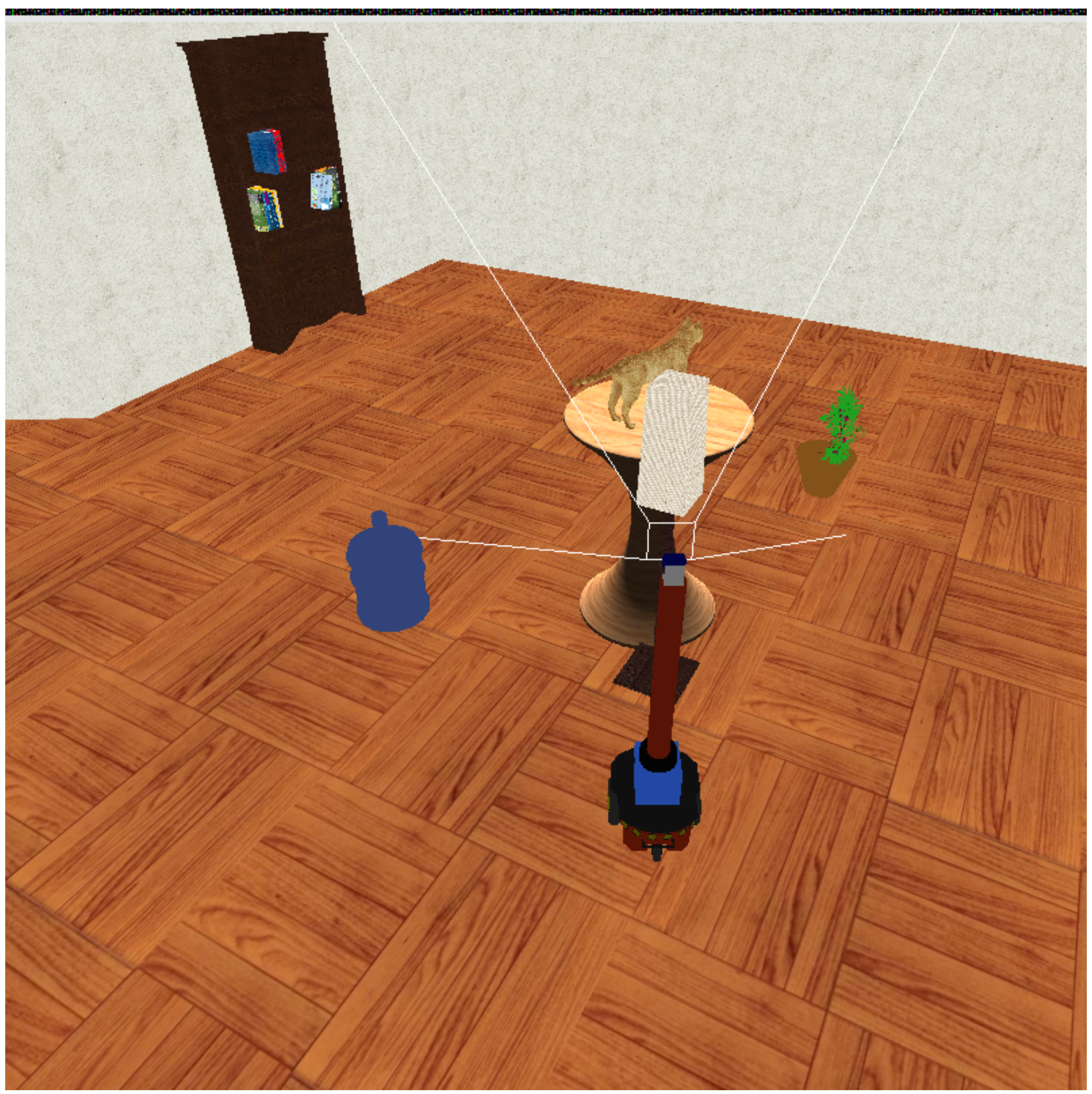
Scene object	σ	# of sensing locations	Path length	Planning time (ms)	% of confirmation
	0.00	14.385	13.305	42.443	100
Cat	0.25	14.550	13.441	46.178	100
	0.50	14.320	13.201	45.284	100
	0.00	10.825	9.122	361.993	100
Bottle	0.25	11.470	9.418	394.074	100
	0.50	11.335	9.442	389.589	100

Dealing with Obstacles generating motion and visibility constraints

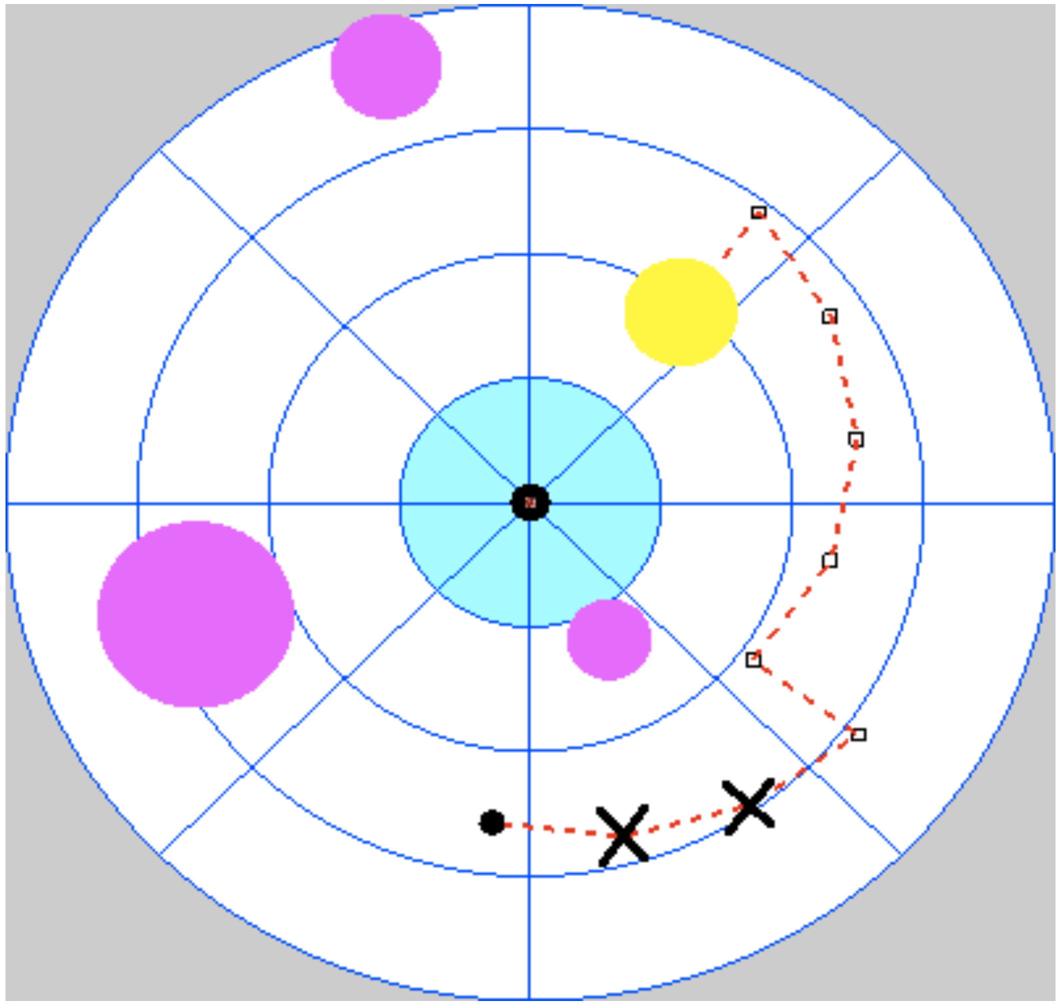




Dealing with Obstacles generating motion and visibility constraints



Dealing with Obstacles generating motion and visibility constraints



[Table 7. Statistics from 100 runs with motion obstacles and the plastic cat as target.

Planning horizon	# of sensing locations	Path length	Planning time (ms)	% of confirmation
3	22.24	20.84	840.13	99

Real-world experiments

- Environment

Real-world experiments

- We use a 24-cell decomposition and a 16-cell decomposition.

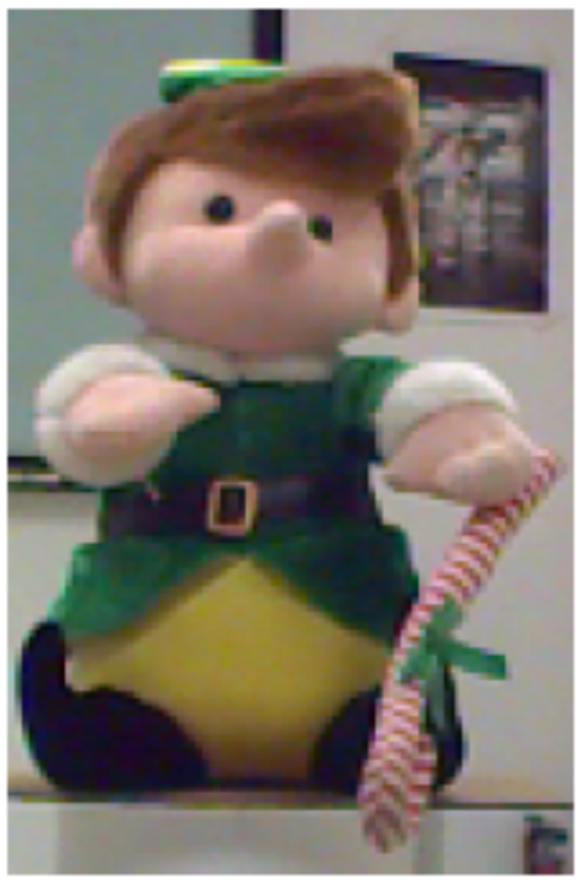
- For each target **T**, the detector DT uses a deformable part model algorithm [2] trained on a set of images taken from a single cell cg of the decomposition.

  [2] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan, "Object Detection with Discriminatively Trained Part Based Models", Trans. on Pattern Analysis and Machine Intelligence, 2010.

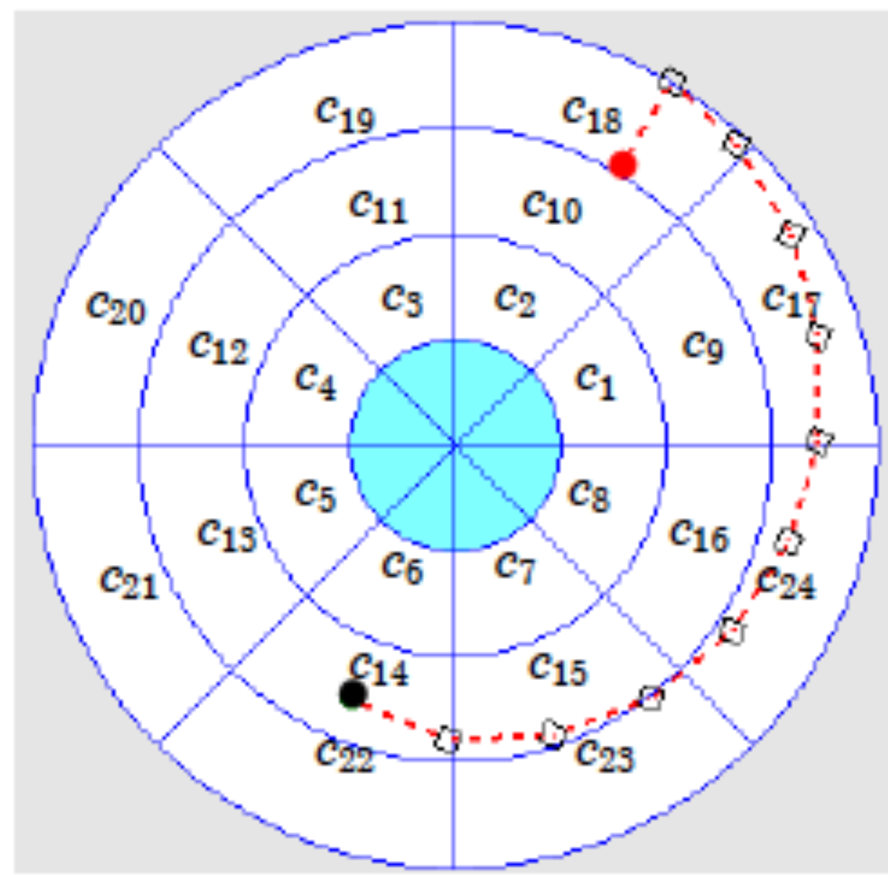
- 6 score values as observation.

Exp #4

- Target



Plush Elf



Exp #4

Exp #4

- Statistics

Planning horizon	# of sensing locations	Path length	Planning time (ms)	% of confirmation
2	14.00	6.685	78.52	100
3	15.00	6.713	1641.39	100
4	14.50	6.851	154880.6	80

Exp #4

- Statistics

Planning horizon	# of sensing locations	Path length	Planning time (ms)	% of confirmation
4	12.00	5.732	119232.2	100

Exp #5

Optimization Criterion

- This approach differs from entropy minimization approaches commonly used in pure localization problems

- Our approach does not minimize entropy by concentrating the probability mass on a particular object (the target) among several other objects as was proposed in [8]

Conclusions

- In this work, we propose an approach to confirm the detection of a given target with a mobile robot equipped with a vision sensor.
- We proposed a strategy mixing robot localisation and target confirmation using the target´s appearance.
- We test our approach in simulations verifying its functionality, and also showing its capability to include distinctive features of the target’s appearance to differentiate the target from similar objects.

Conclusions

- We presented real world experiments testing different targets with different appearances, textures and sizes.
- We also tested different illumination conditions, workspace decomposition sizes and training the detector from more than one training cell.

Future work

- It would be interesting to generate a plan over a larger number of degrees of freedom

Future work

- Experiments wiith a mobile manipulator robot

This work was partially presented in:

- I. Becerra, L. Valentin, R. Murrieta-Cid, and J. C. Latombe. Appearance-based motion strategies for object detection. In Proc. IEEE International Conference on Robotics and Automation, pages 6455–6461, 2014.

To appear in:

- I. Becerra, L. Valentin, R. Murrieta-Cid, and J. C. Latombe. Reliable Confirmation of an Object Identity by a Mobile Robot: A Mixed Appearance/Localization-Driven Motion Approach, Accepted to International Journal of Robotics Research, 2016.

Thanks…

Questions?