A Hierarchical Approach to Sonar-Based Landmark Detection in Mobile Robots

Stefano Rizzi*, Dario Maio, Matteo Golfarelli*****

DEIS - University of Bologna
Viale Risorgimento, 2
40136 Bologna
Italy

Abstract. Landmark recognition in autonomous robots entails extracting symbolic information from sensory data, and is a crucial phase in bridging the gap between the symbolic and the sub-symbolic worlds. In this paper we propose a technique which identifies landmarks by matching a set of templates with a local occupancy grid built by sonar measures; another coarse-grained occupancy grid is used to decide when and where template matching is useful. Our approach is robust since the occupancy grid is a probabilistic model which represents the uncertainty on sensory data; hence, it may work well even on incomplete or noisy landmark shapes. It is fast since, during template matching, the second-level occupancy grid enables the search area to be restricted; thus, the other activities of the robot are not slowed down. The paper reports the results of experimental tests executed on a Pioneer mobile robot equipped with an array of sonars, a compass, an odometer and a camera.

1 Introduction

Several authors agree that navigation in autonomous robots should be accomplished by coupling reactive motion control and capability of planning paths at higher levels of abstraction [4]. Both reactive behaviour and high-level path planning are based on a description of the environment, but each has different demands. The former requires a well-rooted correspondence between entities in the real world and their internal representation, and needs essentially local sensor-based information. On the other hand, too much detail in the description may become overwhelming for abstract path planning which is more easily carried out by adopting a symbolic representation where useless details are neglected.

These requirements cannot be easily met if only one representation formalism is adopted; thus, in [9] we argued that navigation-oriented knowledge of a structured environment should be represented at two levels, symbolic and sub-symbolic, using different formalisms. In the hybrid approach to navigation for autonomous robots we proposed, the *landmarks* in the environment draw a "boundary line" in the environmental knowledge, corresponding to the functional separation between abstract path planning and physical trajectory planning. Paths of landmarks are planned at the symbolic level, on a hierarchy of graphs whose vertices represent landmarks and clusters of landmarks; inter-landmark navigation is achieved by relying on detailed sub-symbolic representations of the routes between pairs of landmarks.

Recognizing landmarks entails extracting symbolic information from sensory data, and is a crucial phase in bridging the gap between the symbolic and the sub-symbolic worlds. The recognition technique depends essentially on which sensor(s) are mounted on the robot, as well as on the semantics which landmarks are required to express. The most common sensors are monocular and binocular cameras [1], lasers and sonars [7]. The semantics associated to landmarks in the literature varies significantly: some approaches define landmarks with low informative content but which can easily be identified; for example, in [6] the authors define a landmark as a place that optimizes a measure of distinctiveness (e.g. differences of the distances from near objects). Others search selectively for landmarks with high distinctive power [2]; using these landmarks improves the recognition reliability but makes identification more complex and less successful.

* Email: srizzi@deis.unibo.it
** Email: dmaio@deis.unibo.it
*** Email: mgolfarelli@deis.unibo.it

From an analysis of the approaches in the literature it is clear that a landmark-identification algorithm should be robust when dealing with the uncertainty due to the errors in the robot self-positioning and in sensor interpretation, and fast in order to avoid slowing down the robot's other activities. The technique we propose in this paper meets these requirements for a robot working in a semi-structured environment, i.e., an environment where landmark shapes and dimensions are known *a priori* but other unknown objects (obstacles) may be present.

In our approach, landmarks are identified by matching a set of templates with a local occupancy grid built by sonar measures; another coarse-grained occupancy grid is used to decide when and where template matching is useful. Our technique is robust since the occupancy grid is a probabilistic model which represents the uncertainty on sensory data; thus, it may work well even on incomplete or noisy landmark shapes. It is fast since, during template matching, the second-level occupancy grid allows the search area to be restricted. We are currently experimenting our approach on a Pioneer I mobile platform (by Real World Interface Inc.) equipped with an array of sonars, a compass, an odometer and a camera.

The paper is organized as follows. In Section 2 the environment representation formalism is outlined. In Section 3 the identification algorithm is described; some experimental results are shown in Section 4.

2 Environment Representation

As stated above, different environment representation formalisms are needed to successfully carry out all the functionalities of an autonomous robot. In particular, our architecture adopts a symbolic graph-based representation for high-level path planning and an analogic grid-based representation to perform other navigational tasks such as obstacle avoidance, landmark detection and inter-landmark navigation. In particular, the analogic description is built by interpreting and combining sensory data; the landmark detection algorithm bridges the gap between the two worlds by defining the nodes of the symbolic description.

2.1 Symbolic Level: The Hierarchic Graph

Symbolic knowledge of the environment is organized on multiple layers at different abstraction levels [9]; each level is structured as a graph whose vertices and arcs represent, respectively, places and connections between them. Within the graph at the lowest abstraction level, vertices correspond to *landmarks* and arcs to *routes* (feasible trajectories between landmarks). Each arc is labeled with the cost paid when covering the corresponding route; each vertex with the position of the corresponding landmark and its category, determined by the template matching algorithm. Within the graphs at the higher levels, each vertex corresponds to a *cluster*, i.e., a connected sub-graph of the graph at the level below (see Figure 1). Clusters are obtained by means of a clustering algorithm based on topological and metric criteria [8].

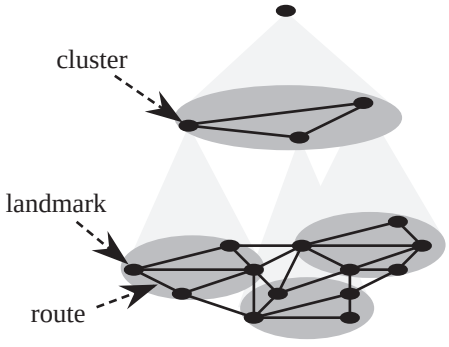


Fig. 1.: Symbolic knowledge representation.

2.2 Analogic Level: The Occupancy Grid

Occupancy grids [3] are multi-dimensional (typically 2-D) tessellations of space; each cell c_{ij} stores the probability of being occupied at a given time, $P(s(c_{ij}) = occ)$. New sensory data acquired modify the stored

values according to the Bayesian law. Let $P(s(c_{ij}) = occ \mid \{r\}_t)$ be the current estimate of the state of a cell c_{ij} based on observation $\{r\}_t = \{r_1, r_2, ...r_t\}$; given the new measurement, r_{t+1}, the improved estimate is given by:

$$P(s(c_{ij}) = occ \mid \{r\}_{t+1}) = \frac{p(r_{t+1} \mid s(c_{ij}) = occ) \cdot P(s(c_{ij}) = occ \mid \{r\}_t)}{\sum_{s(c_{ij})} p(r_{t+1} \mid s(c_{ij})) \cdot P(s(c_{ij}) \mid \{r\}_{t+1})}.$$

In this recursive formulation, the previous estimate of the cell state serves as the prior and is directly obtained from the occupancy grid; $p(r_{t+1} \mid s(c_{ij}) = occ)$ is the probability *a posteriori* that the new observation be r_{t+1}, given $s(c_{ij}) = occ$, and is defined by the sensor model.

Occupancy grids, in this basic formulation, suffer from two problems when used for real-world applications. The first, typically occurring in indoor applications, is due to specular reflections. Specular reflection is a property of active time-of-flight sensors such as sonars and radars, in which the energy from the device is specularly reflected by an angled surface, and may bounce on several surfaces before returning to the device. Specular readings, unlike readings in which the beam is reflected diffusely back to the device, do not give direct information of the distance from the nearest surface.

The second problem is the presence of redundant readings. When a new reading is taken in the same position and orientation (*pose*) of a previous one, it does not introduce further information. For instance, by taking multiple readings from the same pose we cannot detect specular readings.

Both problems can be partially overcome by adopting the MURIEL method to update the occupancy grid [5]. This method manages each sonar reading differently, according to its probability of being specular, and copes with redundant readings by associating each cell with an independence log which stores the pose from which each reading was taken (of course, only one reading from each pose is considered).

The estimate of the state of a cell depends on both $P(s(c_{ij}) = occ \mid \{r\}_{t+1})$ and $P(s(c_{ij}) = free \mid \{r\}_{t+1})$. Consider, for example, a broken sensor that always gives the same reading $r = \delta$, no matter what the environment. In this case it is

$$P(r_{t+1} = \delta \mid s(c_{ij}) = occ) = P(r_{t+1} = \delta \mid s(c_{ij}) = free) = 1$$

so the state of c_{ij} has no effect on the sensors and its probability should not change. This can be obtained by the *log-odds* formulation, where the cell state is defined as:

$$O(s(c_{ij}) = occ) = \frac{P(s(c_{ij}) = occ)}{P(s(c_{ij}) = free)},$$

$$O(s(c_{ij}) = occ|\{r\}_{t+1}) = \frac{P(s(c_{ij}) = occ \mid \{r\}_{t+1})}{P(s(c_{ij}) = free \mid \{r\}_{t+1})}.$$

It should be noted that the value which the MURIEL method associates to each cell of the occupancy grid is not an occupancy probability but rather an occupancy *prevision*, expressed on a logarithmic scale ranging from 0 (absolutely impossible) to $+\infty$ (absolutely true). The *odds-likelihood posterior* can be computed as:

$$\begin{aligned} O(s(c_{ij}) = occ \mid \{r\}_{t+1}) &= \frac{p(\{r\}_{t+1} \mid s(c_{ij}) = occ)}{P(\{r\}_{t+1} \mid s(c_{ij}) = free)} \cdot O(s(c_{ij}) = occ) \\ &\doteq \lambda(\{r\}_{t+1} \mid s(c_{ij}) = occ) \cdot O(s(c_{ij}) = occ). \end{aligned}$$

The occupancy prevision is then linearized and quantized within the range [0..255] (127 means equiprobability of being free or occupied); the resulting value is called OL and stored in the occupancy grid. Linearizing the occupancy prevision, as opposed to using the logarithmic scale, speeds up the decisional process and enables the occupancy previsions for the second level occupancy grid to be computed simply as the weighted average of the occupancy previsions for the first level occupancy grid (see Section 3.1).

When designing an occupancy grid to be used for recognizing landmarks, a crucial parameter is the spatial resolution of cells. In fact, the smaller the side of cells, l, the higher the definition in representing objects, and the higher the computational cost for processing the grid. On the other hand, recognizing landmarks with complex irregular shapes requires high definition.

3 Landmark Recognition

Our approach to landmark recognition consists of three steps; the first and the second are aimed at determining when and where the third, the actual identification step, will be executed. The two filter steps enable the computational cost of the identification process, which is primarily due to the template matching algorithm, to be reduced.

3.1 Evaluation of Sensory Data

The robot moves in the environment following its exploration strategy. Every new sonar pattern read causes the *first level occupancy grid* (OG1) to be updated, and may potentially lead to identifying a landmark. Unfortunately, applying the template matching algorithm whenever OG1 is updated has prohibitive computational costs.

For this reason, we introduce a *second level occupancy grid* (OG2) whose cells correspond to square areas in OG1 including $n \times n$ cells (see Figure 2). The value stored in a cell of OG2 is the weighted average of the

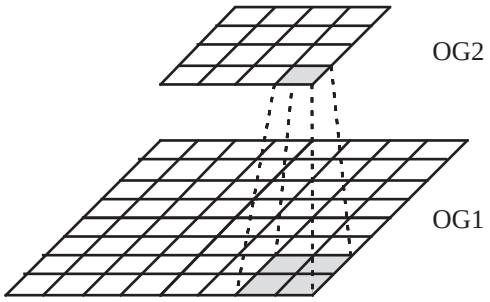


Fig. 2.: Relation between the first and the second level occupancy grid ($n = 2$).

occupancy previsions for all the corresponding cells in OG1. The likelihood OL of cell c^2_{kh} is computed as:

$$OL(s(c^2_{kh}) = occ \mid \{r\}_t) = \frac{\sum_{i=k\cdot n}^{(k+1)n-1} \sum_{i=h\cdot n}^{(h+1)n-1} (w_{ij} \cdot OL(s(c_{ij}) = occ \mid \{r\}_t))}{\sum_{i=k\cdot n}^{(k+1)n-1} \sum_{i=h\cdot n}^{(h+1)n-1} w_{ij}}$$

where

$$w_{ij} = \begin{cases} 1 & \text{if } OL(s(c_{ij}) = occ \mid \{r\}_t) \leq 127 \\ \alpha & \text{otherwise} \end{cases}$$

($\alpha > 1$ is a parameter). The weights w_{ij} are defined in such a way as to give more importance to the occupied cells of OG1. In fact, the only cells whose occupancy prevision is greater than 127 are those lying on the edge of the objects (sonar measures do not enable any occupancy prevision to be made for the cells inside); thus, if a non-weighted average were used, even the OG2 cells corresponding to occupied areas would yield an occupancy prevision lower than 127 (see Figure 3).

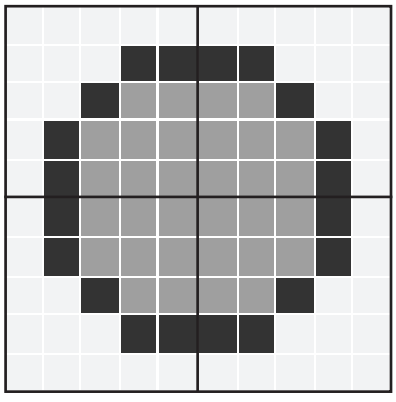


Fig. 3.: Representation of a circular landmark on OG1. White cells indicate free space, black cells indicate the landmark edges. The second level occupancy grid has $n = 5$; it should be noted that for each cell in OG2 the number of free cells is greater than the number of those occupied.

A critical factor in the definition of OG2 is the value given to n, that should depend on both l and d_{min}, the diameter of the smallest landmark. In fact, the area of each cell of OG2 should not overcome that

occupied by the smallest landmark, in order to avoid including free cells of OG1; on the other hand, it should be large enough to significantly reduce the complexity of template matching. We assume that

$$n \leq \lfloor \frac{d_{min}}{l} \rfloor.$$

The value stored in a cell of OG2 represents a prevision of the corresponding area being occupied by an object. Whenever a new sonar reading modifies some cells in OG1, the corresponding cells in OG2 are also updated. The template matching algorithm is launched only when the information acquired is considered to be sufficient to recognize a landmark, that is, when (1) the new occupancy prevision for the updated cell(s) of OG2 is greater than a threshold value and (2) the occupancy prevision has increased significantly with respect to the last time the matching algorithm was executed on the cell(s). Let OG2 be updated by the new reading r_{t+1}; the template matching algorithm is triggered if both inequalities below are satisfied:

- $OL(c^2_{kh} \mid \{r\}_{t+1}) > \tau_{min}$. Threshold τ_{min} expresses the minimum level of certainty below which we may assume that there is no object in the corresponding region of OG1.
- $OL(c^2_{kh} \mid \{r\}_{t+1}) - OL(c^2_{kh} \mid \{r\}_t) > \tau_{var}$. Threshold τ_{var} expresses the minimum required increase in certainty.

3.2 Selection of the Matching Region

When the two above conditions are verified, it is necessary to determine the set of OG1 cells which will be searched by the template matching algorithm; this set is constructed incrementally by using the information contained in OG2.

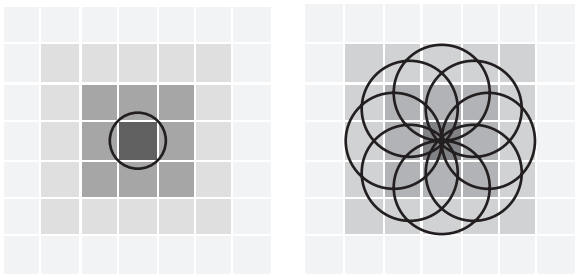
Initially, an area of interest on OG2 is selected by considering the smallest neighborhood of c^2_{kh} which can contain a landmark, i.e., the set of cells c^2_{mn} for which:

$$Distance8(c^2_{kh}, c^2_{mn}) < 1 + \lfloor \frac{d_{min}}{2nl} \rfloor.$$

This inequality is satisfied when the smallest landmark is centered in c^2_{kh}, as shown in Figure 4.a. This set is then iteratively extended with the adjacent cells whose occupancy prevision is greater than a threshold τ_{min}. Iteration is stopped when no adjacent "promising" cells can be found or the area selected stretches to cover the neighborhood determined by the cells c^2_{mn} for which:

$$Distance8(c^2_{kh}, c^2_{mn}) < 1 + \lfloor \frac{d_{max}}{nl} \rfloor$$

(see Figure 4.b).



(a) (b)

Fig. 4.: The smallest (a) and the largest (b) neighborhoods that can be selected for template matching.

Template matching is then executed on all the cells of OG1 corresponding to the cells selected on OG2.

3.3 Template Matching

Within a semi-structured environment, a set $\{g_k\}$ of categories of landmarks corresponding to objects which the robot may meet during navigation is defined a priori. The choice of the categories depends on the nature of the environment. The best candidates are objects delimited by either plane or convex surfaces; in fact, concave surfaces may easily produce sonar reflections.

Recognition is carried out by considering, for each category g_k of landmarks, a set of templates $\{t_{kv}\}$, each reproducing the landmark contour with a different orientation respect to the grid. The number of different orientations depends on the resolution of OG1 and is a trade-off between localization accuracy and efficiency; obviously, it is also determined by the existence of symmetries in the landmark shape.

A template includes only the cells belonging to the edge of the landmark. In fact, although the occupancy prevision for a cell located inside a landmark should not change in time, undesired specular reflections may cause some readings to update its value; on the other hand, the border cells are affected less significantly because the high number of readings makes the occupancy prevision more reliable.

The *matching degree* between a template and the pattern defined by an area of OG1 is calculated by centring the template on a cell of the pattern and comparing the occupancy prevision gradients for each pair of corresponding cells. Gradient values carry more information than occupancy previsions, since they are determined by the cell neighborhoods; they are computed[1] by the Sobel operator ∇. The matching degree between pattern r and template t_{kv} in cell $\mathbf{q} \in r$ is defined as:

$$Match(r, t_{kv}, \mathbf{q}) = \\ = 1 - \frac{\sum_{i,j \in t_{kv}} \sqrt{(\nabla^x(r)[q^x + i, q^y + j] - \nabla^x(t_{kv})[i,j])^2 + (\nabla^y(r)[q^x + i, q^y + j] - \nabla^y(t_{kv})[i,j])^2}}{z_{kv} \cdot \mu}$$

where z_{kv} is the number of cells belonging to t_{kv} and μ is the highest mismatch value of the Sobel operator. The matching degree ranges between 0 (no matching) and 1 (full matching); due to its fuzzy nature, it operates particularly well on distorted or incomplete landmark shapes.

Given a pattern r to classify, the template matching phase consists in computing the matching degree, for every template available, in correspondence to each cell of r; for each template t_{kv}, the cell $\overline{\mathbf{q}}_{kv}$ which maximizes the matching degree is determined. Let $\overline{M}_{kv}$ be the maximum matching degree for template t_{kv}; $\overline{M}_{kv}$ expresses our confidence in ascribing r to template t_{kv}. Pattern r is classified as belonging to the category g_h, if any, which satisfies the following constraints:

$$\exists t_{hw} \mid (\overline{M}_{hw} > \tau_{min}) \wedge \forall t_{mn} (\overline{M}_{hw} > \overline{M}_{mn}).$$

When this constraint is not satisfied by any template, the pattern is not classified as a landmark; this situation usually occurs when:

- The robot has not collected enough sonar readings about that region.
- The pattern does not belong to any of the categories defined as landmarks.
- Some wrong sonar readings determine a strongly corrupted reconstruction.

When a pattern is classified as a landmark, it is added as a new vertex within the symbolic graph, and labelled with its position $\overline{\mathbf{q}}_{hw}$ and its category g_h. Actually, in order to make future identifications more reliable, each landmark is also associated with the set of the templates for which $\overline{M}_{kv} > \tau_{min}$.

4 Experimental Results

This section reports the results of experimental tests performed on the Pioneer mobile robot. The tests have been executed in an office environment where three categories of artificial landmarks had been inserted: circle, square, and triangle; the templates have been defined for each category in order to achieve an orientation precision of 15°.

Figure 5 and Table 1 show a simple example where all the categories of landmarks are present. The good quality of the environment reconstruction allows the landmark to be recognized. As the robot moves on the path, its confidence about the correct landmark category increases. It is remarkable that a satisfactory level of confidence is already obtained when only a partial view of the landmark is available to the robot.

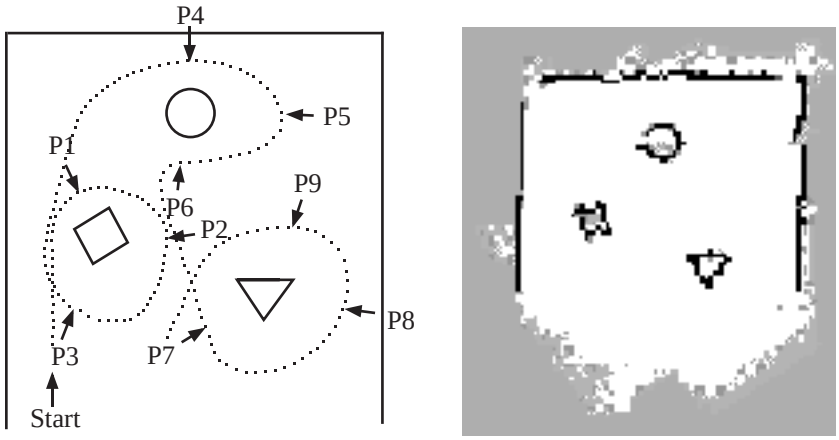


Fig. 5.: First-level occupancy grid (b) built during a simple exploration session (a). Labels from P1 to P9 denote places on the exploration path.

Label	Circle	Triangle	Square
P1	0.2	0.3	0.2
P2	0.4	0.2	0.5
P3	0.4	0.2	0.8
P4	0.2	0.1	0.2
P5	0.4	0.2	0.3
P6	0.7	0.2	0.5
P7	0.1	0.2	0.2
P8	0.2	0.4	0.3
P9	0.2	0.7	0.4

Table 1.: Each row reports the matching degree expressing to what extent the landmark indicated by the arrow in Figure 5.a belongs to the three categories. Template matching was carried out 103 times, each with a mean duration of 0.01 seconds (on a 133 MHz Pentium).

Figure 6 shows how the maximum matching degrees for different templates and categories vary during the recognition of a square landmark.

Several tests have been carried out to detect the limitations of the algorithm. Obviously, most problems arise when wrong sonar readings are taken. In this case, the landmark category may be correctly recognized but its orientation may be missed. Less commonly, the landmark category may be missed. The probability that such errors occur can be reduced by tuning the thresholds. Of course, the thresholds are the result of a trade-off between the risk of incorrect classification and that of completely missing the landmark.

5 Conclusions

In this work we have presented an algorithm for fast landmark identification in semi-structured environments. Our algorithm represents the environment by an occupancy grid and identifies landmarks by matching a set of templates on the occupancy prevision gradient of the cells. A second-level occupancy grid, with coarser granularity, is used to speed up recognition by deciding when and where template matching is needed. Experimental tests executed on the Pioneer mobile robot show that our approach works well even on blurred occupancy grids.

[1] The gradients for all the templates available are computed *a priori*.

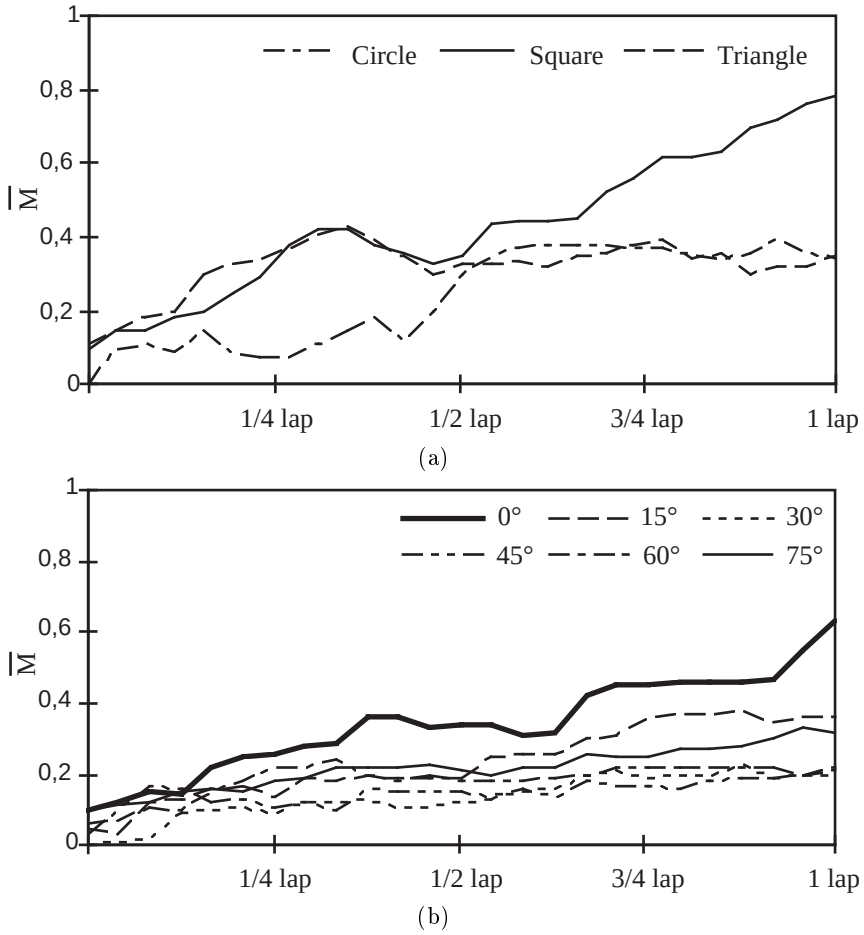


Fig. 6.: Maximum matching degrees returned while the robot is moving around a square-shaped landmark oriented at 0°. Each chart spans one lap made by the robot around the landmark. (a) Different landmark categories; (b) Different templates of the square category.

References

1. H.I. Christensen, N.O. Kirkeby, S. Kristensen, L. Knudsen, and E. Granum. Model-driven vision for in-door navigation. *Robotics and Autonomous Systems*, 12:199–207, 1994.
2. H. Dulimarta and A.K. Jain. Mobile robot localization in indoor environment. *Pattern Recognition*, 30(1):99–111, 1997.
3. A. Elfes. Sonar-based real-world mapping and navigation. *IEEE Journal of Robotics and Automation*, RA-3(3):249–265, 1987.
4. E. Gat. On the role of stored internal state in the control of autonomous mobile robots. *AI Magazine*, 14(1):64–73, 1993.
5. K. Konolige. A refined method for occupancy grid interpretation. In *Proc. Int. Workshop on Reasoning with Uncertainty in Robotics*, Holland, 1995.
6. D. Kortenkamp, M. Huber, and F. Koss et al. Mobile robot exploration and navigation of indoor spaces using sonar and vision. In *Proc. AIAA/NASA Conf. on Intelligent Robots in Field, Factory, Service and Space*, 1994.
7. B.J. Kuipers and Y.T. Byun. A robust, qualitative method for robot spatial learning. In *Proc. AAAI 88*, volume 2, pages 774–779, Saint Paul, Minnesota, 1988.
8. D. Maio, D. Maltoni, and S. Rizzi. Dynamic clustering of maps in autonomous agents. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(11):1080–1091, 1996.
9. D. Maio and S. Rizzi. Layered knowledge architecture for navigation-oriented environment representation. Technical Report 108, CIOC-C.N.R., 1996.

This article was processed using the TEX macro package with SIRS97 style