We are IntechOpen,
the world's leading publisher of
Open Access books
Built by scientists, for scientists

6,900
Open access books available

186,000
International authors and editors

200M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

BOOK CITATION INDEX
CLARIVATE ANALYTICS
INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

Attentional Behaviors for Environment Modeling by a Mobile Robot

Pilar Bachiller, Pablo Bustos and Luis J. Manso
University of Extremadura
Spain

1. Introduction

Building robots capable of interacting in an effective and autonomous way with their environments requires to provide them with the ability to model the world. That is to say, the robot must interpret the environment not as a set of points, but as an organization of more complex structures with human-like meaning. Among the variety of sensory inputs that could be used to equip a robot, vision is one of the most informative ones. Through vision, the robot can analyze the appearance of objects. The use of stereo vision also gives the possibility to extract spatial information of the environment, allowing to determine the structure of the different elements composing it. However, vision suffers from some limitations when it is considered in isolation. On one hand, cameras have a limited field of view that can only be compensated through camera movements. On the other hand, the world is formed by non-convex structures that can only be interpreted by actively exploring the environment. Hence, the robot must move its head and body to give meaning to perceived elements composing its environment.

The combination of stereo vision and active exploration provides a means to model the world. While the robot explores the environment perceived regions can be clustered, forming more complex structures like walls and objects on the floor. Nevertheless, even in simple scenarios with few rooms and obstacles, the robot must be endowed with different abilities to successfully solve the task. For instance, during exploration, the robot must be able to decide where to look at while selecting where to go, avoiding obstacles and detecting what is that it is looking at. From the point of view of perception, there are different visual behaviors that take part in this process, such as those related to look towards what the robot can recognize and model, or those dedicated to maintain itself within safety limits. From the action perspective, the robot has to move in different ways depending on internal states (i.e. the status of the modeling process) and external situations (i.e. obstacles in the way to a target position). Both perception and action should influence each other in such a way that deciding where to look at depends on what the robot is doing, but also in a way that what is being perceived determines what the robot can or can not do.

Our solution to all these questions relies heavily on visual attention. Specifically, the foundation of our proposal is that attention can organize the perceptual and action processes by acting as an intermediary between both of them. The attentional connection allows, on one hand, to drive the perceptual process according to the behavioral requirements and, on the other hand, to modulate actions on the basis of the perceptual results of the attentional control. Thus, attention solves the where to look problem and, additionally, attention prevents

behavioral disorganization by limiting possible actions than can be performed in a given situation. Based on this double functionality, we have developed an attention-based control scheme that generates autonomous behavior in a mobile robot endowed with a 4 dof's (degrees of freedom) stereo vision head. The proposed system is a behavioral architecture that uses attention as the connection between perception and action. Behaviors modulate the attention system according to their particular goals and generate actions consistent with the selected focus of attention. Coordination among behaviors emerges from the attentional nature of the system, so that the robot can simultaneously execute several independent, but cooperative, behaviors to reach complex goals. In this paper, we apply our control architecture to the problem of environment modeling using stereo vision by defining the attentional and behavioral components that provide the robot with the capacity to explore and model the world.

2. Environment modeling using vision

As a first approach towards the environment modeling, we focus on indoor environments composed by several rooms connected through doors. Rooms are considered approximately rectangular and may contain objects on the floor.

During exploration, perceived visual regions are stored in a 3D occupancy grid which constitutes a discrete representation of a certain zone of the environment. This occupancy grid is locally used, so, when the robot gets into a new room, the grid is reseted. Each cell of this grid contains, among other attributes, the certainty degree about the occupancy of the corresponding volume of the environment. The certainty value decreases as the distance to the perceived region increases, assuming this way possible errors in the parametrization of the stereo pair. In addition, the certainty increases as a region is perceived over time in the same position. Thus, stable regions produces higher occupancy values than unstable ones. Cells with a high certainty degree are used for detecting a room model fitting the set of perceived regions. Once the model of the current room can be considered stable, it is stored in an internal representation that maintains topological and metric information of the environment.

Several approaches on mobile robotics propose the use of topological representation to complement the metric information of the environment. In (Thrun, 1998) it is proposed to create off-line topological graphs by partitioning metric maps into regions separated by narrow passages. In (Simhon & Dudek, 1998) the environment is represented by a hybrid topological-metric map composed by a set of local metric maps called *islands of reliability*. (Tomatis et al., 2003) describes the environment using a global topological map that associates places which are metrically represented by infinite lines belonging to the same places. (Van Zwynsvoorde et al., 2000) constructs a topological representation as a route graph using Voronoï diagrams. In (Yan et al., 2006) the environment is represented by a graph whose nodes are crossings (corners or intersections). (Montijano & Sagues, 2009) organizes the information of the environment in a graph of planar regions.

In our approach, the topological representation encodes entities of higher level than the ones mentioned above. Each node of the topological graph represents a room and each edge describes a connection between two rooms. In addition, instead of maintaining a parallel metric map, each topological node contains a minimal set of metric information that allows building a metric map of a place of the environment when it is needed. This approach reduces drastically the amount of computations the robot must perform to maintain an internal representation of the environment. In addition, it can be very helpful for solving certain tasks in an efficient way, such as global navigation or self-localization.

2.1 Room modeling

Since rooms are assumed to be rectangular and its walls perpendicular to the floor, the problem of modeling a room from a set of regions can be treated as a rectangle detection problem. Several rectangle detection techniques can be found in the literature (Lagunovsky & Ablameyko, 1999; Lin & Nevatia, 1998; Tao et al., 2002). Most of them are based on a search in the 2D point space (for instance, a search in the edge representation of an image) using line primitives. These methods are computationally expensive and can be very sensitive to noisy data. In order to solve the modeling problem in an efficient way, we propose a new rectangle detection technique based on a search in the parameter space using a variation of the Hough Transform (Duda & Hart, 1972; Rosenfeld, 1969).

For line detection, several variations of the Hough Transform have been proposed (Matas et al., 2000; Palmer et al., 1994). The extension of the Hough Transform for rectangle detection is not new. (Zhu et al., 2003) proposes a *Rectangular Hough Transform* used to detect the center and orientation of a rectangle with known dimensions. (Jung & Schramm, 2004) proposes a *Windowed Hough Transform* that consists of searching rectangle patterns in the Hough space of every window of suitable dimensions of an image.

Our approach for rectangle detection uses a 3D version of the Hough Transform that facilitates the detection of segments instead of lines. This allows considering only those points that belong to the contour of a rectangle in the detection process. The Hough space is parameterized by (θ, d, p), being θ and d the parameters of the line representation ($d = x\cos(\theta) + y\sin(\theta)$) and $|p|$ the length of a segment in the line. For computing p it is assumed that one of the extreme points of its associated segment is initially fixed and situated at a distance of 0 to the perpendicular line passing through the origin. Under this assumption, being (x, y) the other extreme point of the segment, its *signed* length p can be computed as:

$$p = x\cos(\theta + \pi/2) + y\sin(\theta + \pi/2) \tag{1}$$

Using this representation, any point (x, y) contributes to those points (θ, d, p) in the Hough space that verifies:

$$d = x\cos(\theta) + y\sin(\theta) \tag{2}$$

and

$$p >= x\cos(\theta + \pi/2) + y\sin(\theta + \pi/2) \tag{3}$$

Equation 2 represents every line intersecting the point as in the original Hough Transform. The additional condition expressed by equation 3 limits the point contribution to those line segments containing the point. This allows computing the total number of points included in a given segment. For instance, given a segment with extreme points $V_i = (x_i, y_i)$ and $V_j = (x_j, y_j)$ and being H the 3D Hough space, the number of points that belong to the segment, which is denoted as $H_{i\leftrightarrow j}$, can be computed as:

$$H_{i\leftrightarrow j} = |H(\theta_{i\leftrightarrow j}, d_{i\leftrightarrow j}, p_i) - H(\theta_{i\leftrightarrow j}, d_{i\leftrightarrow j}, p_j)| \tag{4}$$

where $\theta_{i\leftrightarrow j}$ and $d_{i\leftrightarrow j}$ are the parameters of the common line to both points and p_i and p_j the *signed* lengths of the two segments with non-fixed extreme points V_i and V_j, respectively, according to equation 1.

Since a rectangle is composed by four segments, the 3D Hough space parameterized by (θ, d, p) allows computing the total number of points included in the contour of the rectangle. Thus, considering a rectangle expressed by its four vertices $V_1 = (x_1, y_1)$, $V_2 = (x_2, y_2)$, $V_3 = (x_3, y_3)$

and $V_4 = (x_4, y_4)$ (see figure 1), the number of points of its contour, denoted as H_r, can be computed as:

$$H_r = H_{1\leftrightarrow2} + H_{2\leftrightarrow3} + H_{3\leftrightarrow4} + H_{4\leftrightarrow1} \tag{5}$$

Considering the restrictions about the segments of the rectangle and using the equation 4, each $H_{i\leftrightarrow j}$ of the expression 5 can be rewritten as follows:

$$H_{1\leftrightarrow2} = |H(\alpha, d_{1\leftrightarrow2}, d_{4\leftrightarrow1}) - H(\alpha, d_{1\leftrightarrow2}, d_{2\leftrightarrow3})| \tag{6}$$

$$H_{2\leftrightarrow3} = |H(\alpha + \pi/2, d_{2\leftrightarrow3}, d_{1\leftrightarrow2}) - H(\alpha + \pi/2, d_{2\leftrightarrow3}, d_{3\leftrightarrow4})| \tag{7}$$

$$H_{3\leftrightarrow4} = |H(\alpha, d_{3\leftrightarrow4}, d_{2\leftrightarrow3}) - H(\alpha, d_{3\leftrightarrow4}, d_{4\leftrightarrow1})| \tag{8}$$

$$H_{4\leftrightarrow1} = |H(\alpha + \pi/2, d_{4\leftrightarrow1}, d_{3\leftrightarrow4}) - H(\alpha + \pi/2, d_{4\leftrightarrow1}, d_{1\leftrightarrow2})| \tag{9}$$

being α the orientation of the rectangle as expressed in figure 1 and $d_{i\leftrightarrow j}$ the normal distance of the origin to the straight line defined by the points V_i and V_j.

Since H_r expresses the number of points in a rectangle r defined by $(\alpha, d_{1\leftrightarrow2}, d_{2\leftrightarrow3}, d_{3\leftrightarrow4}, d_{4\leftrightarrow1})$, the problem of obtaining the best rectangle given a set of points can be solved by finding the combination of $(\alpha, d_{1\leftrightarrow2}, d_{2\leftrightarrow3}, d_{3\leftrightarrow4}, d_{4\leftrightarrow1})$ that maximizes H_r. This parametrization of the rectangle can be transformed into a more practical representation defined by the five-tuple (α, x_c, y_c, w, h), being (x_c, y_c) the central point of the rectangle and w and h its dimensions. This transformation can be achieved using the following expressions:

$$x_c = \frac{d_{1\leftrightarrow2} + d_{3\leftrightarrow4}}{2} \cos(\alpha) - \frac{d_{2\leftrightarrow3} + d_{4\leftrightarrow1}}{2} \sin(\alpha) \tag{10}$$

$$y_c = \frac{d_{1\leftrightarrow2} + d_{3\leftrightarrow4}}{2} \sin(\alpha) + \frac{d_{2\leftrightarrow3} + d_{4\leftrightarrow1}}{2} \cos(\alpha) \tag{11}$$

$$w = d_{2\leftrightarrow3} - d_{4\leftrightarrow1} \tag{12}$$

$$h = d_{3\leftrightarrow4} - d_{1\leftrightarrow2} \tag{13}$$

In order to compute H_r, the parameter space H is discretized assuming the rank $[-\pi/2, \pi/2]$ for θ and $[d_{min}, d_{max}]$ for d and p, being d_{min} and d_{max} the minimum and maximum distance, respectively, between a line and the origin. The sampling step of each parameter is chosen according to the required accuracy. Figure 1 shows an example of rectangle representation in the discretized parameter space. Each pair of parallel segments of the rectangle is represented in the corresponding orientation plane of the discrete Hough space: $H(\alpha_d)$ for one pair of segments and $H((\alpha + \pi/2)_d)$ for the other one, being α_d and $(\alpha + \pi/2)_d$ the discrete values associated to α (the rectangle orientation) and $(\alpha + \pi/2)$, respectively. For each orientation plane, it is represented how many points contribute to each cell (d_d, p_d), i.e. how many points belong to every segment of the corresponding orientation. A high histogram contribution is represented in the figure with a dark gray level, while a low contribution is depicted with an almost white color. As it can be observed, the maximum contributions are found in parallel segments with displacements of w_d and h_d, which are the discrete values associated to the rectangle dimensions.
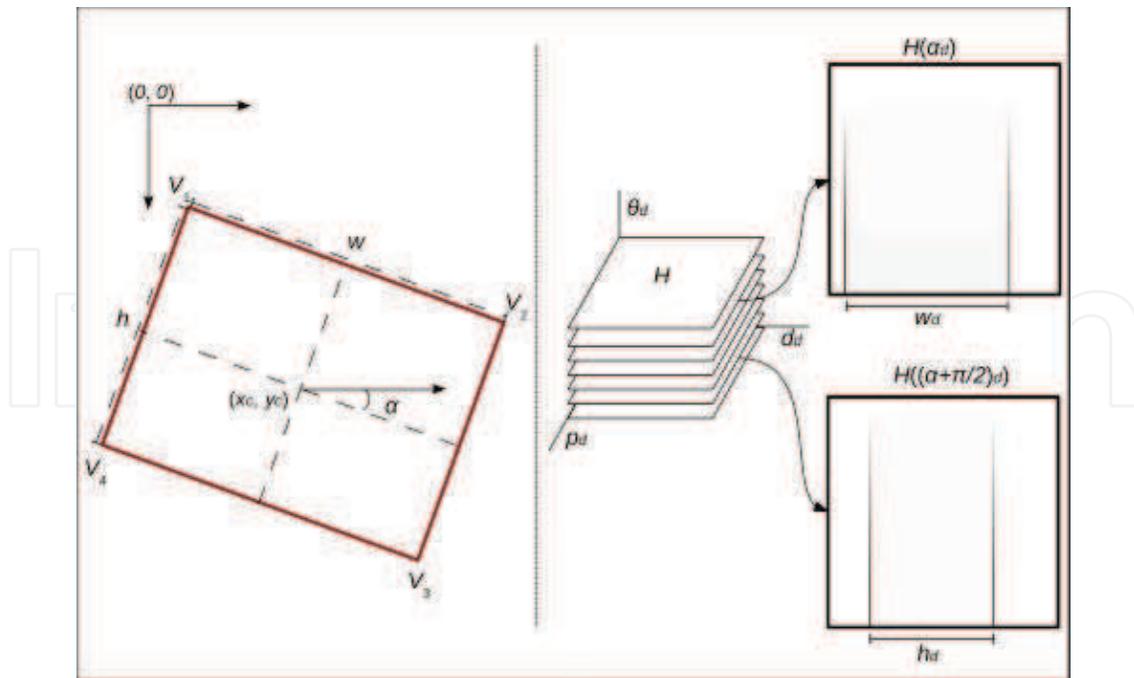
Fig. 1. Rectangle detection using the proposed 3D variation of the Hough Transform (see the text for further explanation)

This rectangle detection technique is used to obtain a room model that fits the points stored in the 3D occupancy grid. Walls are considered to have a maximum height and, therefore, only points situated at a certain rank of height in the grid are used for detecting the model. Assuming that this rank is in the interval $[0, Z_{wall}]$ and being G the 3D occupancy grid and τ the minimum occupancy value to considered a non empty region of the environment, the proposed method for room modeling can be summarized in the following steps:

1. Initialize all the cells of the discrete Hough space H to 0.

2. For each cell, $G(x_d, y_d, z_d)$, such that $G(x_d, y_d, z_d).occupancy > \tau$ and $z_d \epsilon [0, Z_{wall}]$:

   Compute the real coordinates (x, y) associated to the cell indexes (x_d, y_d).

   For $\theta_d = \theta_{dMin} \ldots \theta_{dMax}$:
   (a) Compute the real value θ associated to θ_d.
   (b) Compute $d = x \cos(\theta) + y \sin(\theta)$.
   (c) Compute the discrete value d_d associated to d.
   (d) Compute $p = x \cos(\theta + \pi/2) + y \sin(\theta + \pi/2)$.
   (e) Compute the discrete value p_d associated to p.
   (f) For $p'_d = p_d \ldots d_{dMax}$: increment $H(\theta_d, d_d, p'_d)$ by 1.

3. Compute $\underset{\alpha, d_{1 \leftrightarrow 2}, d_{2 \leftrightarrow 3}, d_{3 \leftrightarrow 4}, d_{4 \leftrightarrow 1}}{\arg \max} H_r(\alpha, d_{1 \leftrightarrow 2}, d_{2 \leftrightarrow 3}, d_{3 \leftrightarrow 4}, d_{4 \leftrightarrow 1})$.

4. Obtain the rectangle $r = (\alpha, x_c, y_c, w, h)$ using equations 10, 11, 12 and 13.

2.2 Door detection

Doors are free passage zones that connect two different rooms, so they can be considered as empty segments of the corresponding room rectangle (i.e. segments without points). Taking this into account, once the room model is obtained, doors can be detected by analyzing each

wall segment in the 3D Hough space. Thus, for each segment of the rectangle defined by V_i and V_j, two points $D_k = (x_k, y_k)$ and $D_l = (x_l, y_l)$ situated on the inside of that segment constitutes a door segment if it is verified:

$$H_{k \leftrightarrow l} = |H(\theta_{i \leftrightarrow j}, d_{i \leftrightarrow j}, p_k) - H(\theta_{i \leftrightarrow j}, d_{i \leftrightarrow j}, p_l)| = 0 \tag{14}$$

being $\theta_{i \leftrightarrow j}$ and $d_{i \leftrightarrow j}$ the parameters of the straight line defined by V_i and V_j and p_k and p_l the *signed* lengths of the segments for D_k and D_l:

$$p_k = x_k \cos(\theta_{i \leftrightarrow j} + \pi/2) + y_k \sin(\theta_{i \leftrightarrow j} + \pi/2) \tag{15}$$

$$p_l = x_l \cos(\theta_{i \leftrightarrow j} + \pi/2) + y_l \sin(\theta_{i \leftrightarrow j} + \pi/2) \tag{16}$$

Assuming $p_i \leq p_k < p_l \leq p_j$ and a minimum length l for each door segment, the door detection process can be carried out by verifying equation 14 for every pair of points between V_i and V_j, such that $p_l - p_k \geq l$. Starting from the discrete representation of the Hough space, this process can be summarized in the following steps:

1. Compute the discrete value θ_d associated to θ_{i-j}.

2. Compute the discrete value d_d associated to d_{i-j}.

3. Compute the discrete value p_{di} associated to p_i.

4. Compute the discrete value p_{dj} associated to p_j.

5. Compute the discrete value l_d associated to l (minimum length of doors).

6. $p_{dk} \leftarrow p_{di}$

7. While $p_{dk} \leq p_{dj} - l_d$:
    (a) $p_{dl} \leftarrow p_{dk} + 1$
    (b) While $p_{dl} < p_{dj}$ and $|H(\theta_d, d_d, p_{dk}) - H(\theta_d, d_d, p_{dl})| = 0$: $p_{dl} \leftarrow p_{dl} + 1$
    (c) If $p_{dl} - p_{dk} > l_d$:
        i. Compute the real value p_k associated to p_{dk}.
        ii. Compute the real value p_l associated to $(p_{dl} - 1)$.
        iii. Compute the door limits D_k and D_l from p_k and p_l.
        iv. Insert the new door segment with extreme points D_k and D_l to the list of doors.
    (d) $p_{dk} \leftarrow p_{dl}$

The output of this method is the list of doors of the wall segment delimited by the vertices V_i and V_j. Each door is represented by its extreme points D_k and D_l, which are computed in step 7.(c).iii. from p_k and p_l. Since both points verify the line equation ($d_{i \leftrightarrow j} = x \cos(\theta_{i \leftrightarrow j}) + y \sin(\theta_{i \leftrightarrow j})$), using 15 and 16, their coordinates can be computed as follows:

$$x_k = d_{i \leftrightarrow j} \cos(\theta_{i \leftrightarrow j}) - p_k \sin(\theta_{i \leftrightarrow j}) \tag{17}$$

$$y_k = d_{i \leftrightarrow j} \sin(\theta_{i \leftrightarrow j}) + p_k \cos(\theta_{i \leftrightarrow j}) \tag{18}$$

$$x_l = d_{i \leftrightarrow j} \cos(\theta_{i \leftrightarrow j}) - p_l \sin(\theta_{i \leftrightarrow j}) \tag{19}$$

$$y_l = d_{i \leftrightarrow j} \sin(\theta_{i \leftrightarrow j}) + p_l \cos(\theta_{i \leftrightarrow j}) \tag{20}$$

2.3 Topological and metric representation of the environment

The detected rooms and doors are modeled and used to build a topological representation of the environment. In this representation, the environment is described as an undirected graph whose vertices represent the different explored rooms (see figure 2). An edge linking two vertices expresses the existence of a door that connects two rooms. This is a very useful representation for the robot to effectively move around man-made environments. For instance, the robot could analyze the graph to obtain the minimum path connecting any two rooms. Moreover, this representation can be extended using recursive descriptions to express more complex world structures like buildings. Thus, a building could be represented by a node containing several interconnected subgraphs. Each subgraph would represent a floor of the building and contain a description of the interconnections between the different rooms and corridors in it.


(a)   (b)
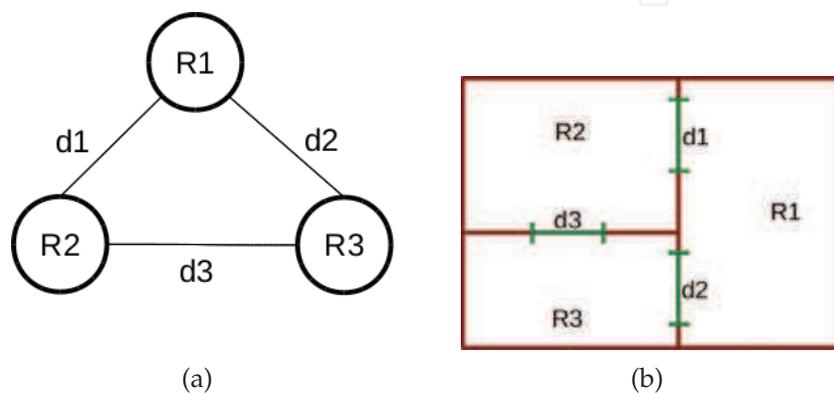
Fig. 2. Topological representation (a) of an environment (b) composed by three intercommunicated rooms.

Using this topological graph, a minimal set of metric information is maintained. Specifically, each vertex of the graph stores the parametric description of the corresponding room and its doors. Metric maps of the environment can be easily computed from this representation when necessary (e.g. when the robot has to compute a metric path connecting two rooms). In order to maintain this basic metric representation, each room model contains a reference frame (F_r) which expresses the location of the room in relation to a global reference frame (F_w). The room reference frame is located at the room center with a rotation given by the room orientation. Thus, being $r = (\alpha, x_c, y_c, w, h)$ the rectangle that models a given room, the transformation matrix (T_r) that relates F_r with F_w is defined as:

$$T_r = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 & x_c \\ \sin(\alpha) & \cos(\alpha) & 0 & y_c \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{21}$$

This matrix provides the transformation $p_w = T_r p_r$, being p_w and p_r the homogeneous coordinates of a 3D point viewed from F_w and F_r, respectively. In the same way, the coordinates of a point in a room $r1$ can be transformed into coordinates expressed in other room ($r2$) reference frame by applying the corresponding sequence of transformations:

$$p_{r2} = T_{r2}^{-1} T_{r1} p_{r1} \tag{22}$$

where p_{r1} is a point situated inside the room $r1$, p_{r2} is the same point viewed from the reference frame of the room $r2$ and T_{r1} and T_{r2} are the transformation matrices of the reference frames of $r1$ and $r2$, respectively.

If two rooms, $r1$ and $r2$, are communicated by a door, points of the door are common to both rooms. Assume a door point d_{r1} viewed from the room $r1$ and the corresponding point d_{r2} of the room $r2$. The metric representation of both rooms would ideally be subject to the following restriction:

$$\|T_{r2}d_{r2} - T_{r1}d_{r1}\|^2 = 0 \tag{23}$$

(a)   (b)

(c)   (d)
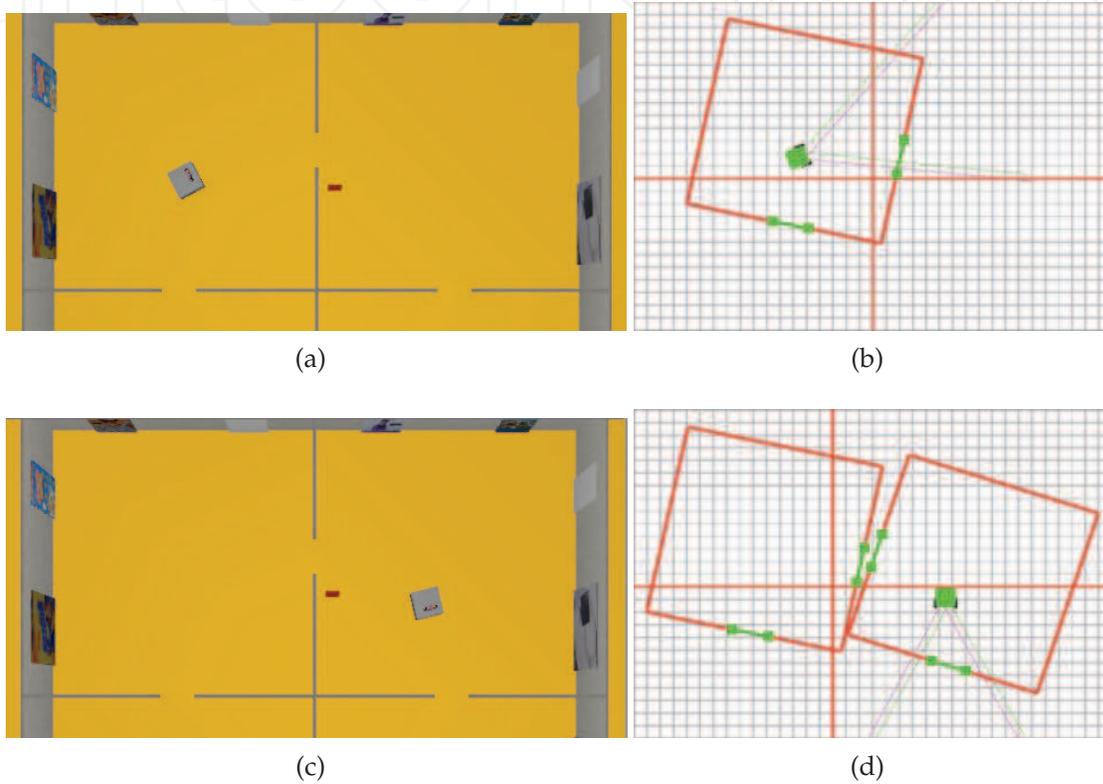
Fig. 3. Deviation between the position of a door common to two rooms caused by odometric errors (results obtained using the gazebo simulator): (a) overhead view of the scene after the creation of the first room model; (b) metric representation of the first room; (c) overhead view of the scene after the creation of the second room model, (d) metric representation of the two rooms.

During exploration, odometric errors cause deviations between the positions of a common door to two adjacent rooms and therefore expression 23 is usually not verified when a new room model is created (see figure 3). However, these deviations allow computing how the reference frame of each room model should be modified in order to fulfill with the *common door restriction*. Thus, given $d^{(1)}$ and $d^{(2)}$, extreme points of the common door, the rotational and translational deviations ($\triangle \alpha$ and $\triangle t$) between two adjacent room models can be computed as:

$$\triangle \alpha = \arctan\left(\frac{y_{r2}^{(1)} - y_{r2}^{(2)}}{x_{r2}^{(1)} - x_{r2}^{(2)}}\right) - \arctan\left(\frac{y_{r1}^{(1)} - y_{r1}^{(2)}}{x_{r1}^{(1)} - x_{r1}^{(2)}}\right) \tag{24}$$

$$\triangle t = d_{r2}^{(1)} - \begin{pmatrix} \cos(\triangle\alpha) & -\sin(\triangle\alpha) & 0 \\ \sin(\triangle\alpha) & \cos(\triangle\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} d_{r1}^{(1)} \tag{25}$$

being $d_{rj}^{(i)} = (x_{rj}^{(i)}, y_{rj}^{(i)}, z_{rj}^{(i)})^T$ the door point $d^{(i)}$ expressed in the reference frame of the model rj.

Assuming an egocentric metric system, the robot pose and the reference frame of the current room model are not affected by these deviations. This implies that remaining room models must be moved according to equations 24 and 25. Thus, for every room model $ri = (\alpha_i, x_{ci}, y_{ci}, w_i, h_i)$ different than the current one, its reference frame is updated as follows:

$$\alpha_i \leftarrow \alpha_i + \triangle\alpha \tag{26}$$

$$\begin{pmatrix} x_{ci} \\ y_{ci} \\ 0 \end{pmatrix} \leftarrow \begin{pmatrix} \cos(\triangle\alpha) & -\sin(\triangle\alpha) & 0 \\ \sin(\triangle\alpha) & \cos(\triangle\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_{ci} \\ y_{ci} \\ 0 \end{pmatrix} + \triangle t \tag{27}$$

Figure 4 shows the result of applying this correction to the metric representation of figure 3. In case of using an eccentric representation, the robot pose and the reference frame of the current model are corrected applying the deviations $\triangle\alpha$ and $\triangle t$ in inverse order.
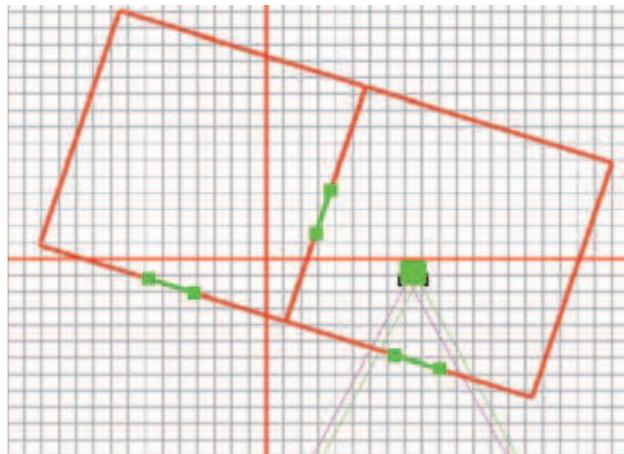

Fig. 4. Metric correction of the environment representation of figure 3 based on the common door restriction.

A similar correction must be carried out to deal with odometric errors when the robot is in a previously modeled room. In such cases, the pose of the robot relative to the room where it is located can be computed according to the new location of the room. To estimate the new location, perceived regions must be used to detect a room model with known dimensions following the detection process of section 2.1. Being $r(i) = (\alpha(i), x_c(i), y_c(i), w, h)$ the room model at instant i and $r(i + 1) = (\alpha(i + 1), x_c(i + 1), y_c(i + 1), w, h)$ a new estimation of the room model at $i + 1$, the model deviation can be computed as:

$$\triangle\alpha = \alpha(i + 1) - \alpha(i) \tag{28}$$

$$\triangle t = \begin{pmatrix} x_c(i + 1) \\ y_c(i + 1) \\ 0 \end{pmatrix} - \begin{pmatrix} \cos(\triangle\alpha) & -\sin(\triangle\alpha) & 0 \\ \sin(\triangle\alpha) & \cos(\triangle\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_c(i) \\ y_c(i) \\ 0 \end{pmatrix} \tag{29}$$

Using these equations, the robot pose in an eccentric representation or the reference frames of room models in an egocentric one are corrected. Figure 5 shows an example. This correction is only applied when there exists no ambiguity in the result of the new estimation of the room model. This means that if the set of perceived regions can be associated to more than one model, the new estimation is rejected.


(a)   (b)

(c)   (d)
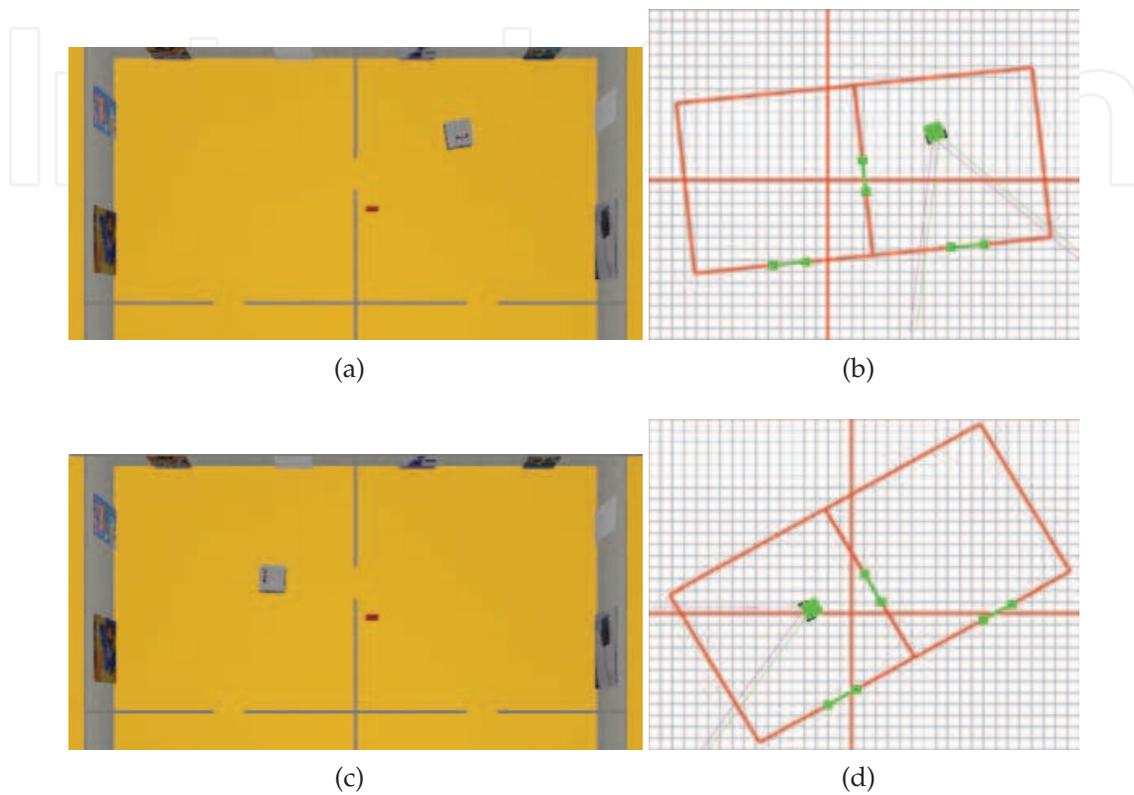
Fig. 5. Metric correction through room model re-estimation.

(a)   (b)
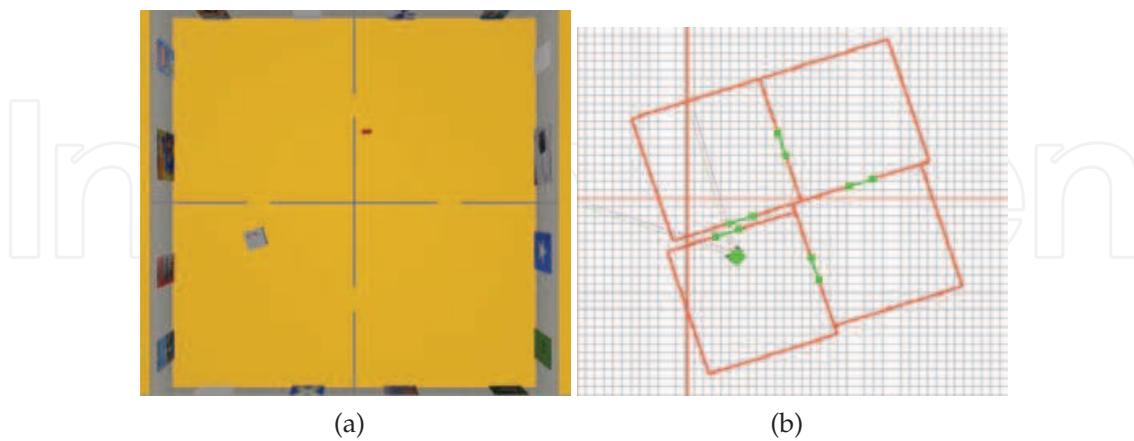
Fig. 6. Metric errors in a loop closing.

Another critical problem to take into account in the creation of a metric representation of the environment are loop closings. The term *loop closing* refers to the return to a previously visited place after an exploration of arbitrary length. These situations occur when the robot detects a new door which is connected to a previously visited room. In such cases, new

corrections must be done to minimize the error in the position of the detected common door (see figure 6). However, this error is caused by an imperfect estimation of the parameters of rooms and doors and, therefore, a unique correction will surely not solve the problem. A solution to this problem is to distribute the parameter adjustment among every model in the metric representation (Olson, 2008). The basic idea of this approach is to minimize a global error function defined over the whole metric representation by introducing small variations in the different elements composing that representation. These variations are constrained by the uncertainty of the measurement, so *high-confident* parameters remain almost unchanged during the error minimization process.

In our environment representation, the global error is defined in terms of deviations between the positions of the doors connecting adjacent rooms. Thus, the error function to minimize can be expressed as:

$$\xi = \sum_{\forall connected(d_{ri}^{(n)}, d_{rj}^{(m)})} \|T_{ri} d_{ri}^{(n)} - T_{rj} d_{rj}^{(m)}\|^2 \tag{30}$$

being $d_{ri}^{(n)}$ and $d_{rj}^{(m)}$ the middle points of a common door expressed in the reference frames of rooms ri and rj, respectively, and T_{ri} and T_{rj} the transformation matrices of such rooms.

To minimize ξ, we use the Stochastic Gradient Descent (Robbins & Monro, 1951), which has proven to be an efficient method to solve similar problems (Olson, 2008). In SGD, the error function is iteratively reduced by randomly selecting a parameter and modifying it using a gradient descent step. Each step is modulated by a *learning rate* λ which is reduced over time to avoid local minima. Being S the set of parameters and ξ the error function, the method proceeds as follows:

1. Initialize λ

2. While not converge:
   (a) randomly select a parameter s_i of the set S
   (b) Compute the step of s_i ($\triangle s_i$) in the gradient descent direction of ξ according to the uncertainty of s_i.
   (c) Modify s_i: $s_i \leftarrow s_i + \lambda \triangle s_i$
   (d) Update λ: $\lambda \leftarrow \lambda/(\lambda+1)$

The set of parameters affecting the error function of equation 30 are those related to room and door models. In door models, the only parameter susceptible to variations is its central position relative to the corresponding wall. Assuming the order of walls in figure 7(a), this position ($c_{rk}^{(l)}$) corresponds to the first coordinate for doors in walls 1 and 3 or to the second one for doors in walls 2 and 4. Thus, an adjustment of a door position $d_{rk}^{(l)}$ through a variation $\triangle c_{rk}^{(l)}$ can be written as:

$$d_{rk}^{(l)} \leftarrow (c_{rk}^{(l)} + \triangle c_{rk}^{(l)}, -h_{rk}/2, 0)^T \quad for\ doors\ in\ wall\ 1 \tag{31}$$

$$d_{rk}^{(l)} \leftarrow (w_{rk}/2, c_{rk}^{(l)} + \triangle c_{rk}^{(l)}, 0)^T \quad for\ doors\ in\ wall\ 2 \tag{32}$$

$$d_{rk}^{(l)} \leftarrow (c_{rk}^{(l)} + \triangle c_{rk}^{(l)}, h_{rk}/2, 0)^T \quad for\ doors\ in\ wall\ 3 \tag{33}$$

$$d_{rk}^{(l)} \leftarrow (-w_{rk}/2, c_{rk}^{(l)} + \triangle c_{rk}^{(l)}, 0)^T \quad for\ doors\ in\ wall\ 4 \tag{34}$$

being h_{rk} and w_{rk} the height and width of the room rk.

Regarding room parameters, potential errors in the detection process may affect only to the estimation of the room dimensions (h_{rk} and w_{rk}). Thus, any variation in a room model should be associated to these parameters. However, since every wall corresponds to a segment of the room model and the uncertainty in the detection process is associated to segments and not to model dimensions, the position of every wall is individually adjusted (see 7(b)). These wall adjustments modify the dimensions of the rooms as follows:

$$h_{rk} \leftarrow h_{rk} + \triangle h_{rk}^{(2)} - \triangle h_{rk}^{(1)} \tag{35}$$

$$w_{rk} \leftarrow w_{rk} + \triangle w_{rk}^{(2)} - \triangle w_{rk}^{(1)} \tag{36}$$

(a)   (b)
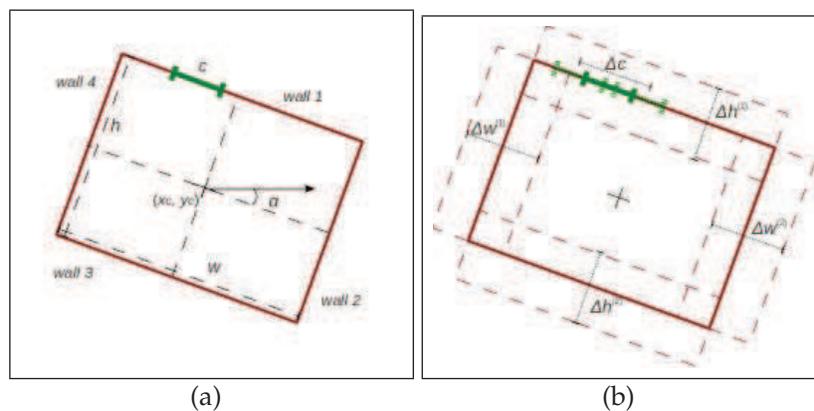
Fig. 7. 7(a) Parametrization of room and door models ; 7(b) Adjustable parameters of rooms and doors.

In addition, changes in wall positions can modify the center of the model. Therefore, for any change in a wall position, the room transformation matrix must be updated as follows:

$$T_{rk} \leftarrow T_{rk} \begin{pmatrix} 1 & 0 & 0 & (\triangle w_{rk}^{(1)} + \triangle w_{rk}^{(2)})/2 \\ 0 & 1 & 0 & (\triangle h_{rk}^{(1)} + \triangle h_{rk}^{(2)})/2 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \tag{37}$$

Thus, using the set of parameters formed by each door central position ($c_{rk}^{(l)}$) and each room wall position ($h_{rk}^{(1)}, h_{rk}^{(2)}, w_{rk}^{(1)}, w_{rk}^{(2)}$), the error function ξ of equation 30 is minimized following the SGD method previously described. It must be taken into account that when the selected parameter is a wall position the transformation matrix of the corresponding room must be updated according to equation 37. Figure 8 shows the result of applying this method to the metric representation of figure 6.

3. The attention-based control architecture

As it was stated in the introduction of this chapter, in order to provide the robot with the ability to explore and model its environment in an autonomous way, it is necessary to endow it with different perceptual behaviors. Perception should be strongly linked to the robot actions in such a way that deciding where to look is influenced by what the robot is doing and also

Fig. 8. Metric correction of the loop closing errors of figure 6.

in a way that the actions of the robot are limited by what is being perceived. These double link between perception and action is solved using the attention-based control architecture (Bachiller et al., 2008).

The proposed attention-based control architecture is composed by three intercommunicated subsystems: the behavioral system, the visual attention system and the motor control system. The behavioral system generates high-level actions that allows keeping different behavioral objectives in the robot. The visual attention system contains the ocular fixation mechanisms that provide the selection and foveatization of visual targets. These two systems are connected to the motor control system, which is responsible of executing the motor responses generated by both of them.

Each high-level behavior modulates the visual attention system in an specific way to get the most suitable flow of visual information according to its behavioral goals. At every execution cycle, the attention system selects a single visual target and sends it to the behavioral system, which executes the most appropriate action according to the received visual information. Thus, the attention system also modulates the behavioral one. This double modulation (from the behavioral system to the attention system and from the attention system to the behavioral one) endows the robot with both deliberative and reactive abilities since it can drive the perceptual process according to the needs or intentions of the robot, but its actions are conditioned by the outside world. This makes the robot interact in an effective way with a real and non-structured environment.

3.1 Attentional requirements

The function of attention in our system is strongly linked to the selection for action mechanisms (Allport, 1987), since it is used to select the most suitable stimulus for action execution. From this point of view, the attention system should maintain the following performance requirements: a) the selection of a visual target should be conditioned by its visual properties; b) this selection should also be influenced by the behavioral intentions or necessities of the robot; c) the system must provide a single focus of attention acting as the only visual input of every high-level behavior; d) the system should be able to simultaneously maintain several visual targets in order to alternate among them covering this way the perceptual needs of every high-level behavior. All these requirements can be fulfilled combining four kinds of attention:

- Bottom-up attention

- Top-down attention

- Overt attention

- Covert attention

These modes of attention can be found at some degree in the different computational models proposed in the literature. Some of them employ a pure bottom-up strategy (Itti & Koch, 2000; Koch & Ullman, 1985), while others integrate contextual information (Torralba et al., 2006) or knowledge about relevant properties of the visual targets along with the sensory data (Frintrop, 2006; Navalpakkam & Itti, 2006). All of them provide overt or covert attention, since the result of the attentional process is the selection of the most salient or conspicuous visual region. If this selection implies eyes movements, it is called overt attention. If it only produces a mental focusing on the selected region, the attention is covert.

Despite the variety of proposals, all these models are characterized by a common aspect: attention control is centralized. It is to say, the result of every processing unit of the system in these models is used by an unique control component that is responsible for driving attention. The centralization of the attentional control presents some problems that prevent from solving key aspects of our proposal. These problems can be summarized in the following three points:

1. Specification of multiple targets

2. Attentional shifts among different targets

3. Reaction to unexpected stimuli

From the point of view of complex actions, the robot needs to maintain several behavioral goals which will be frequently guided by different visual targets. If attentional control is centralized, the specification of multiple visual targets could not work well because the system has to integrate all the selection criteria in an unique support (a saliency or conspicuity map) that represents the relevance of every visual region. This integration becomes complicated (or even unfeasible) when some aspects of one target are in contradiction with the specification of other target, leading sometimes to a wrong attentional behavior. Even though an effective integration of multiple targets could be achieved, another question remains: how to shift attention at the required frequency from one type of target to another one? In a centralized control system, mechanisms as inhibition of return do not solve this question, since the integration of multiple stimuli cancels the possibility of distinguishing among different kinds of targets. A potential solution to both problems could consist of dynamically modulating the visual system for attending only one kind of target at a time. This allows shifting attention among different visual regions at the desired frequency, avoiding any problem related to the integration of multiple targets. However, this solution presents an important weakness: attention can only be programmed to focus on expected things and so the robot could not be able to react to unforeseen stimuli.

In order to overcome these limitations, we propose a distributed system of visual attention, in which the selection of the focus of attention is accomplished by multiple control units called *attentional selectors*. Each attentional selector drives attention from different top-down specifications to focus on different types of visual targets. At any given time, overt attention is driven by one attentional selector, while the rest attends covertly to their corresponding targets. The frequency at which an attentional selector operates overtly is modulated by the high level behavioral units depending on its information requirements. This approach solves the problems described previously. Firstly, it admits the coexistence of different types of visual targets, providing a clearer and simpler design of the selection mechanisms than a centralized approach. Secondly, each attentional selector is modulated to focus attention

on the corresponding target at a given frequency. This prevents from constantly centering attention on the same visual target and guarantees an appropriate distribution of the attention time among the different targets. Lastly, since several attentional selectors can operate simultaneously, covert attention on a visual region can be transformed into overt attention as soon as it is necessary, allowing the robot to appropriately react to any situation.

3.2 A distributed system of visual attention

The proposed visual attention system presents the general structure of figure 9. The perception components are related to image acquisition, detection of regions of interest (Harris-Laplace regions) and extraction of geometrical and appearance features of each detected region. These features are used by a set of components, called attentional selectors, to drive attention according to certain top-down behavioral specifications. Attentional control is not centralized, but distributed among several attentional selectors. Each of them makes its own selection process to focus on an specific type of visual region. For this purpose, they individually compute a saliency map that represents the relevancy of each region according to their top-down specifications. This saliency map acts as a control surface whose maxima match with candidate visual regions to get the focus of attention.

The simultaneous execution of multiple attentional selectors requires including an overt-attention controller that decides which individually selected region gains the overt focus of attention at each moment. Attentional selectors attend covertly to their selected regions. They request the overt-attention controller to take overt control of attention at a certain frequency that is modulated by high-level behavioral units. This frequency depends on the information requirements of the corresponding behavior, so, at any moment, several target selectors could try to get the overt control of attention. To deal with this situation, the overt-attention controller maintains a time stamp for each active attentional selector that indicates when to yield control to that individual selector. Every so often, the overt-attention controller analyses the time stamp of every attentional selector. The selector with the oldest mark is then chosen for driving the overt control of attention. If several selectors share the oldest time stamp, the one with the highest frequency acquires motor control. Frequencies of individual selectors can be interpreted as alerting levels that allow keeping a higher or lower attention degree on the corresponding target. This way, the described strategy gives priority to those selectors with the highest alerting level that require faster control responses.
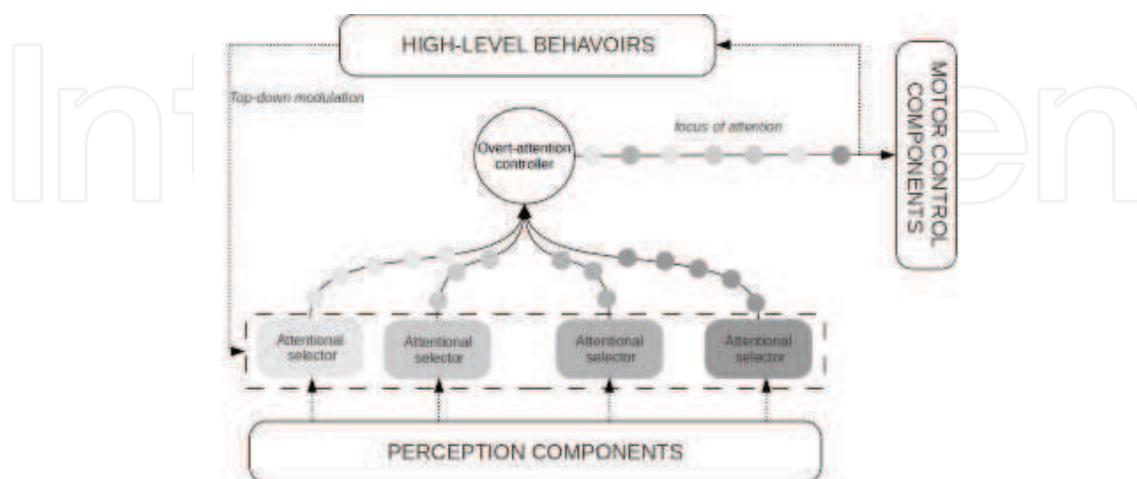

Fig. 9. General structure of the proposed distributed system of visual attention

Once the overt focus of attention is selected, it is sent to the high-level behavioral components. Only actions compatible with the focus of attention are then executed, providing a mechanism of coordination among behaviors. In addition, the selected visual region is centered in the images of the stereo pair, achieving a binocular overt fixation of the current target until another visual target is selected.

Our proposal for this binocular fixation is to use a cooperative control scheme in which each camera plays a different role in the global control. Thus, the 3D fixation movement is separated into two movements: a monocular tracking movement in one on the cameras (the dominant camera) and an asymmetric vergence movement in the other one (secondary camera). This separation allows the saccade that provides the initial fixation on the target to be programmed for a single camera while maintaining a stable focus in both cameras (Enright, 1998). In addition, this scheme provides an effective response to situations in which it is not possible to obtain a complete correspondence of the target in the pair of images due to the change of perspective, the partial occlusion in one of the views or even the non-visibility of the target from one of the cameras.

4. Active modeling using the attention-based control architecture

In an active modeling process, the robot must be able to explore and build a representation of the environment in an unsupervised way (i.e. without human intervention). The different perceptive and high-level behaviors taking part in this process have to deal with different issues: look for the walls of the room, detect obstacles in the path when the robot is moving, decide whether what appears to be door is an actual door or not, or decide when to start exploring another place of the environment, among others. To endow the robot with the required capability to solve all these questions, we propose the behavioral system of figure 10 which follows our attention based-control approach.

Each behavioral and attentional component of the proposed system has an specific role in the modeling task. The Active Modeler behavior starts the task by gaining access to the visual information around the robot. For this purpose, it activates an attentional selector, which attends to visual regions of interest situated in front of the robot, and starts turning the robot base around. The rotational velocity varies according to the attentional response in such a way that the speed increases if no visual region is perceived in front of the robot. Once the robot returns to its initial orientation, a first model of the room is obtained. This model is the rectangular configuration of walls that best fits the set of perceived regions. The resulting model is then improved by the Room Verifier behavior, which forces the robot to approach to those regions with higher uncertainty. To accomplish its goal, Room Verifier activates an attentional selector that changes the gaze so the cameras point towards those visual regions situated at high uncertainty zones. At the same time, it sends the goal positions to the Go to Point behavior in order to make the robot approach those zones. This last behavior is connected to an attentional selector of obstacles, which shifts attention towards regions situated close to the trajectory between the robot and the goal position. The Go to Point behavior interprets the incoming visual information as the nearest regions that could interfere in the approach to the destination position and generates changes in the robot trajectory according to this. In this situation, the focus of attention alternates between uncertainty regions and potential obstacles, keeping overt control on one of the targets and covert control on the other one. This behavior forces the robot to react appropriately to obstacles while the goal position can be quickly recovered and updated. Once the whole uncertainty of the model is low enough, it is considered stable and new behaviors take place. Specifically, the robot tries to locate untextured zones on the walls and hypothesizes them as potential doors that
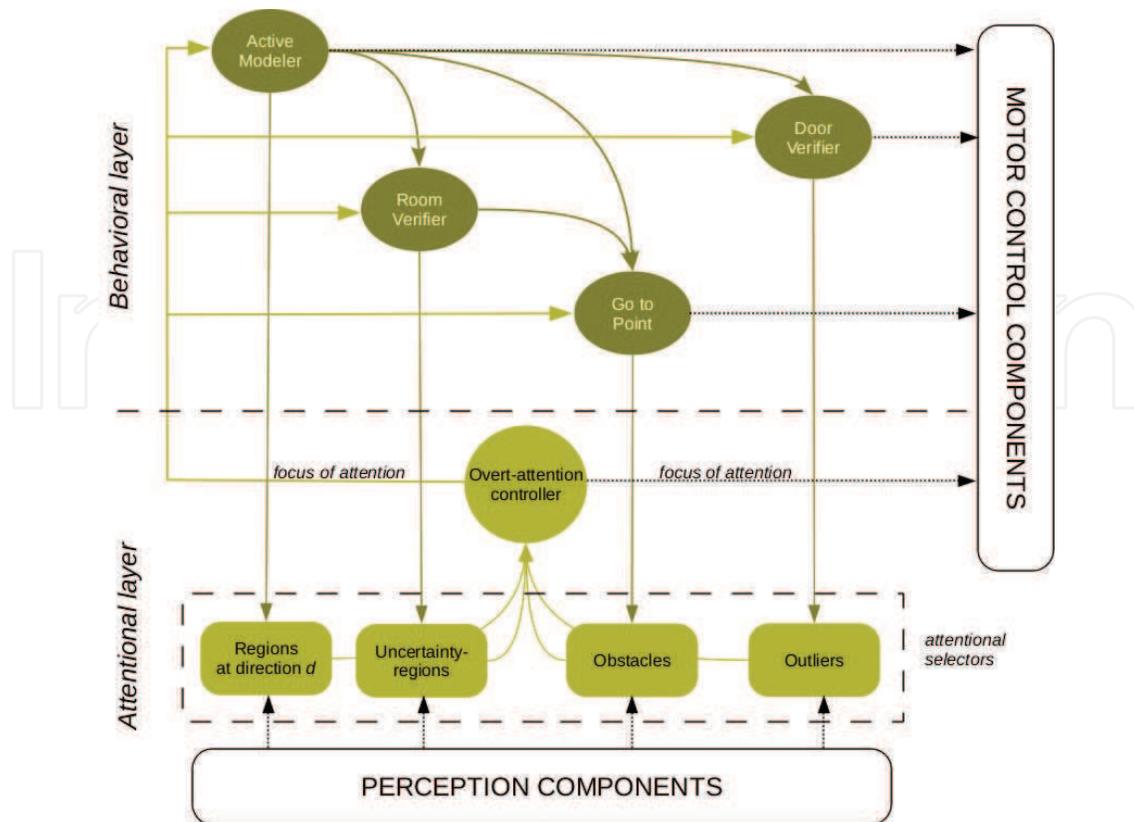
Fig. 10. Scheme of the proposed behavioral system for environment modeling

must be verified. At that time, the Active Modeler behavior reconfigures the corresponding attentional selector to focus on visual regions in the direction of the door where the robot has to approach. Once again, the approaching action is solved by the Go to Point component, taking advantages of the distributed attentional control. When the robot is near enough a potential door, the Door Verifier behavior tries to verify the current hypothesis searching for regions situated behind the corresponding wall. To do so, it activates an attentional selector which selects regions that can be considered as outliers of the room model. If necessary, the Door Verifier behavior spins the base of the robot to cover the whole space of the door. Using the visual information that has been recovered during this process, the robot decides whether to reject or accept the door hypothesis as a real door. If accepted as a real door, the robot gets out of the room to explore unknown spaces driven by the Active Modeler behavior. The room left behind is reduced to a minimal geometric parametrization and stored as a node in the graph representing the topological space. Finally, the whole process is re-initiated and new models are obtained, giving rise over time to a complete topological representation of the environment.

5. Experimental results

We have carried out different experiments in real indoor environments. They were designed to demonstrate the modeling ability of a mobile robot equipped with a stereo vision head. In particular, we used RobEx (figure 11), an open-hardware robotic platform. For the experiments, RobEx was equipped with a 4-dof stereo vision head providing: a neck-movement followed by a common tilt and two independent camera pan movements (Mateos et al., 2010).
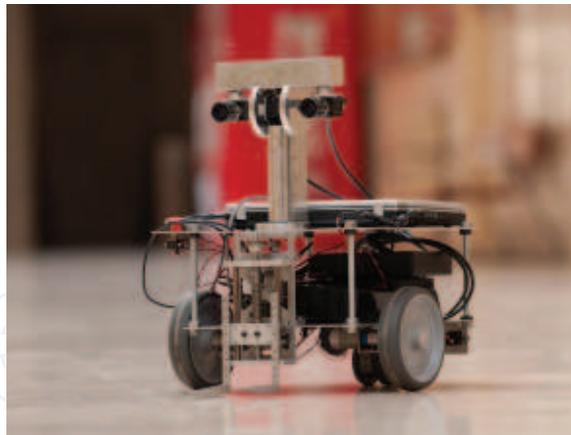
Fig. 11. The RobEx platform.

In the first experiment, we tested the ability of the robot to model a room with several objects on the floor. Figure 12 shows the result of this experiment for the modeling of the room at the right side of figure 13. Figure 12(a) shows the set of perceived regions after an autonomous exploration of the room. These regions are associated to cells of the occupancy grid with high certainty degree. From this set of points, the models of the room and the door are detected (figure 12(b)). The metric representation of the resulting models is shown in figure 12(c). Using this room model, the occupancy grid is updated to cover only the local space inside the room. In addition, the positions of the points near to walls are adjusted according to the detected model. Figure 12(d) shows the result of this adjustment.

In the second experiment, the robot modeled an environment composed by two rooms communicated by a door (figure 14). Figures 15 and 16 show the evolution of the modeling process during the environment exploration. Specifically, for the modeling of the first room, the results of the robot behavior can be observed in figures 15(a) to 15(d): (a) an initial model is created fixating attention on frontal visual regions while the robot base turns around; (b) the robot verifies the room model by keeping attention on high uncertainty zones while approaching them. This allows correcting the model parameters as well as discarding the false door on the right of figure (a); (c) the door is verified and its parameters are adjusted by fixating attention on regions situated outside of the room; (d) the robot gets out of the room to start exploring the second room. Once the robot gets into the new room, the modeling process is repeated for detecting a new model as it is shown in figures 16(a) to 16(c): (a) initial model creation; (b) room model verifying stage; (c) door verifying stage. Finally (figure 16(d)) the deviation between the two room models is corrected applying the common door restriction.

The last experiment tested the ability of the robot to model a more complex configuration of the environment. The environment of the experiment was composed by three intercommunicated rooms with different dimensions (figure 18). Figure 17 shows the modeling results during the exploration of each room. After creating the final model of every room, the deviation of the new model is corrected according to the door connecting the room to an existing model. Finally, it is obtained an environment representation that corresponds to a great extent with the real scene. The error in the dimensions of the first room (see figure 17(e)) is due to the discretization of the Hough space. This kind of errors are considered in the representation of the model as part of its uncertainty. It is important to note that this uncertainty is local to the model. Thus, as it is shown in figure 17, there is no error propagation in the creation of remaining models composing the final representation of the environment.
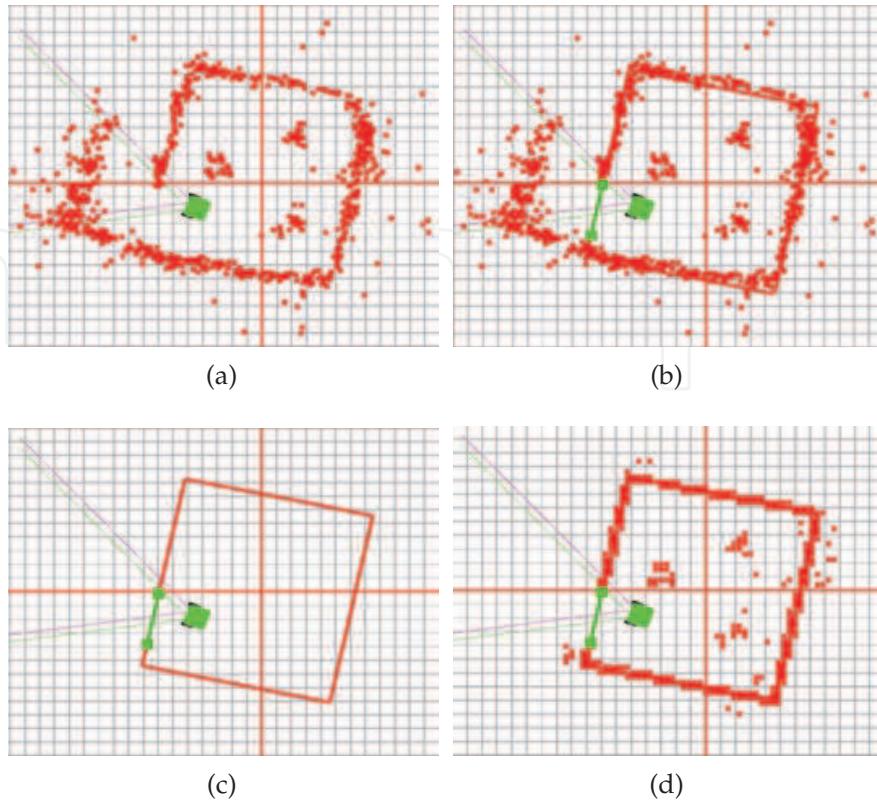
Fig. 12. Room modeling process: (a) 2D view of the perceived regions during the exploration; (b) room and door models detection from the set of points in (a); (c) metric representation of the room; (d) state of the occupancy grid after fixating the room model in (c). These results correspond to the environment shown in 13.
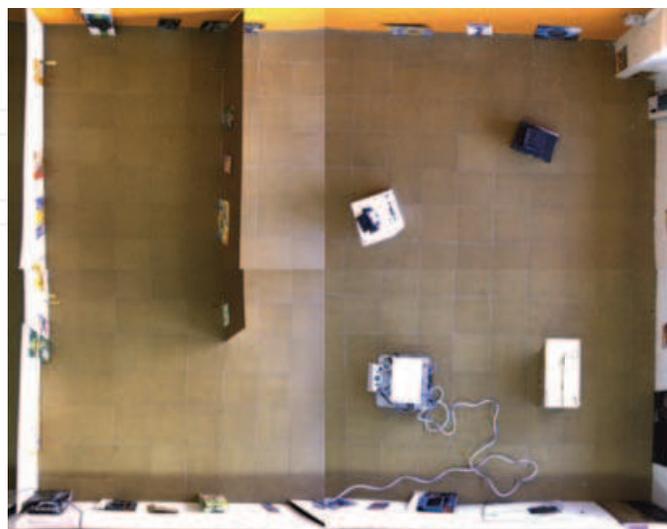
Fig. 13. Overhead view of the real scene of the experiment of figure 12.

Fig. 14. Overhead view of the environment of the experiment of figures 15 and 16.

(a)   (b)

(c)   (d)
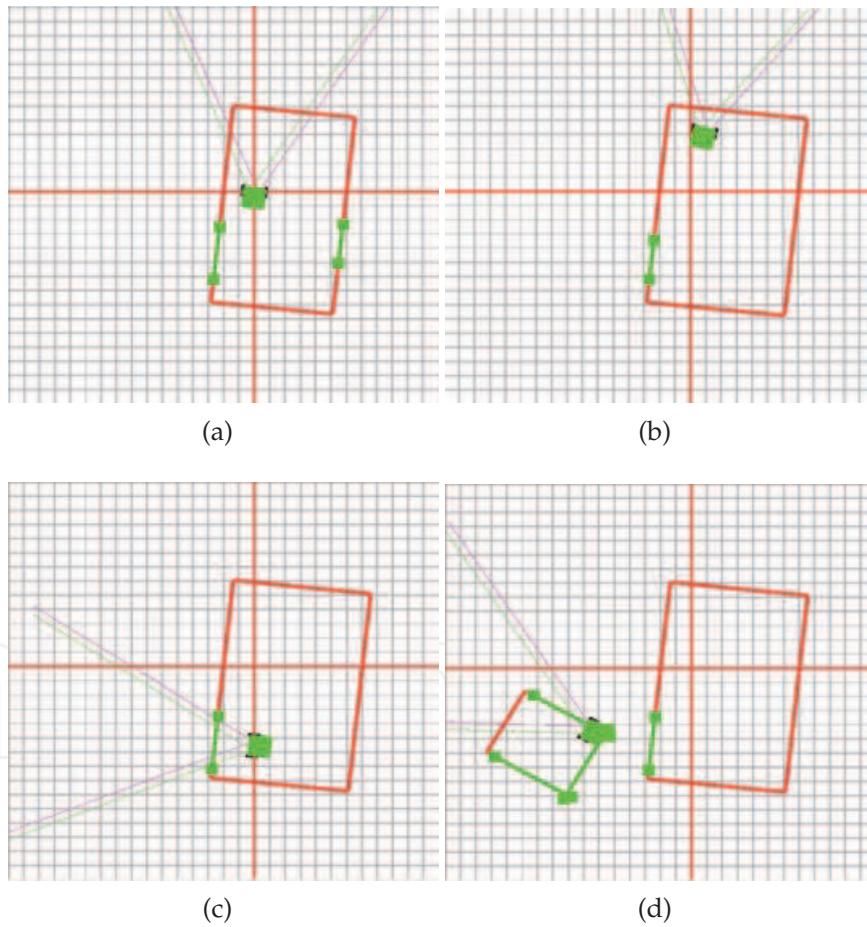
Fig. 15. Modeling of the first room during the exploration of the environment of figure 14.

6. Conclusions and future work

In this chapter, we have presented a behavioral architecture for mobile robots endowed with stereo vision that provides them with the ability to actively explore and model their
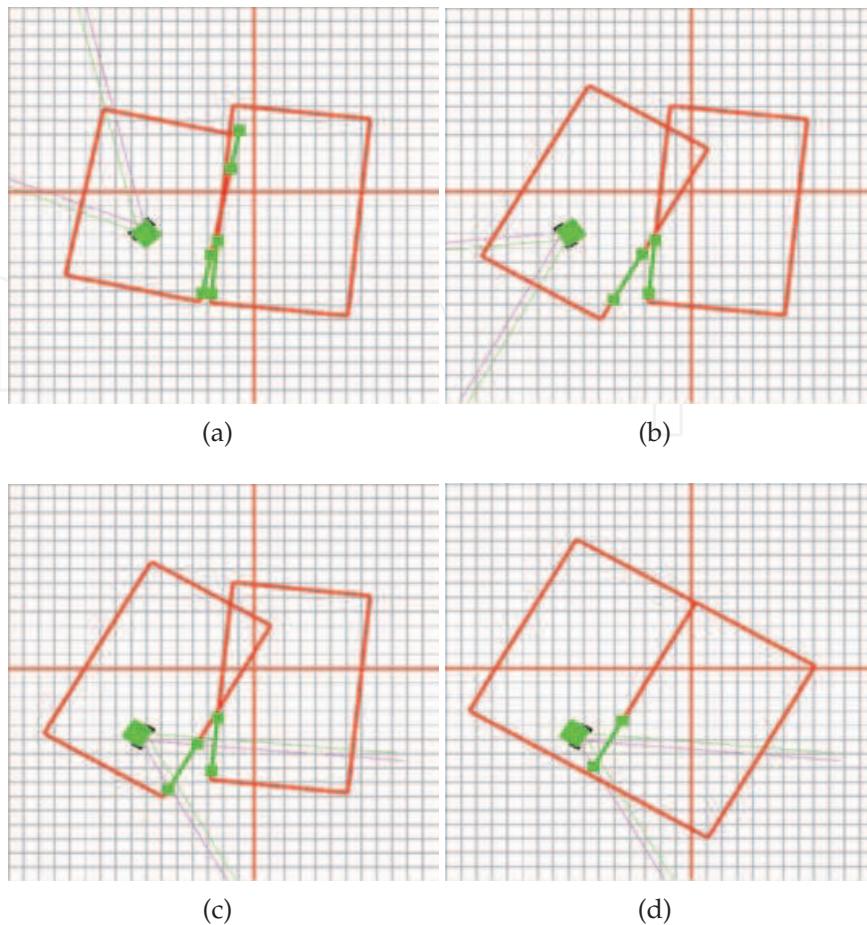
Fig. 16. Modeling of the second room during the exploration of the environment of figure 14.

environments. As an initial approach, it is assumed that the environment is formed by rectangular rooms communicated by doors and may contain objects on the floor. The result of the modeling process is a topological graph that represents the set of rooms (nodes) and their connections (edges). Each node in this representation contains the metric description of the corresponding room model. Using this basic metric information robots do not need to maintain in parallel a metric map of the environment. Instead, this metric map can be built whenever it is necessary from the topological representation. Rooms are modeled using a variation of the Hough Transform which detects segments instead of lines. Deviations between room models caused by odometric errors are easily detected and corrected using the geometric restrictions provided by the door connecting them. In addition, we have proposed methods for robot pose estimation as well as for global metric adjustment in loop closings.

The set of perceptual and high-level behaviors needed to solve the active modeling problem are organized according to our attention-based control architecture. In this architecture, attention is conceived as an intermediary between visual perception and action control, solving two fundamental behavioral questions for the robot: *where to look* and *what to do*. Using this scheme, we have defined the different attentional and high-level behaviors that allow the robot to solve the modeling task in an autonomous way. The resulting behavioral system has been tested in real indoor environments of different complexity. These results prove the effectiveness of our proposal in real scenarios.
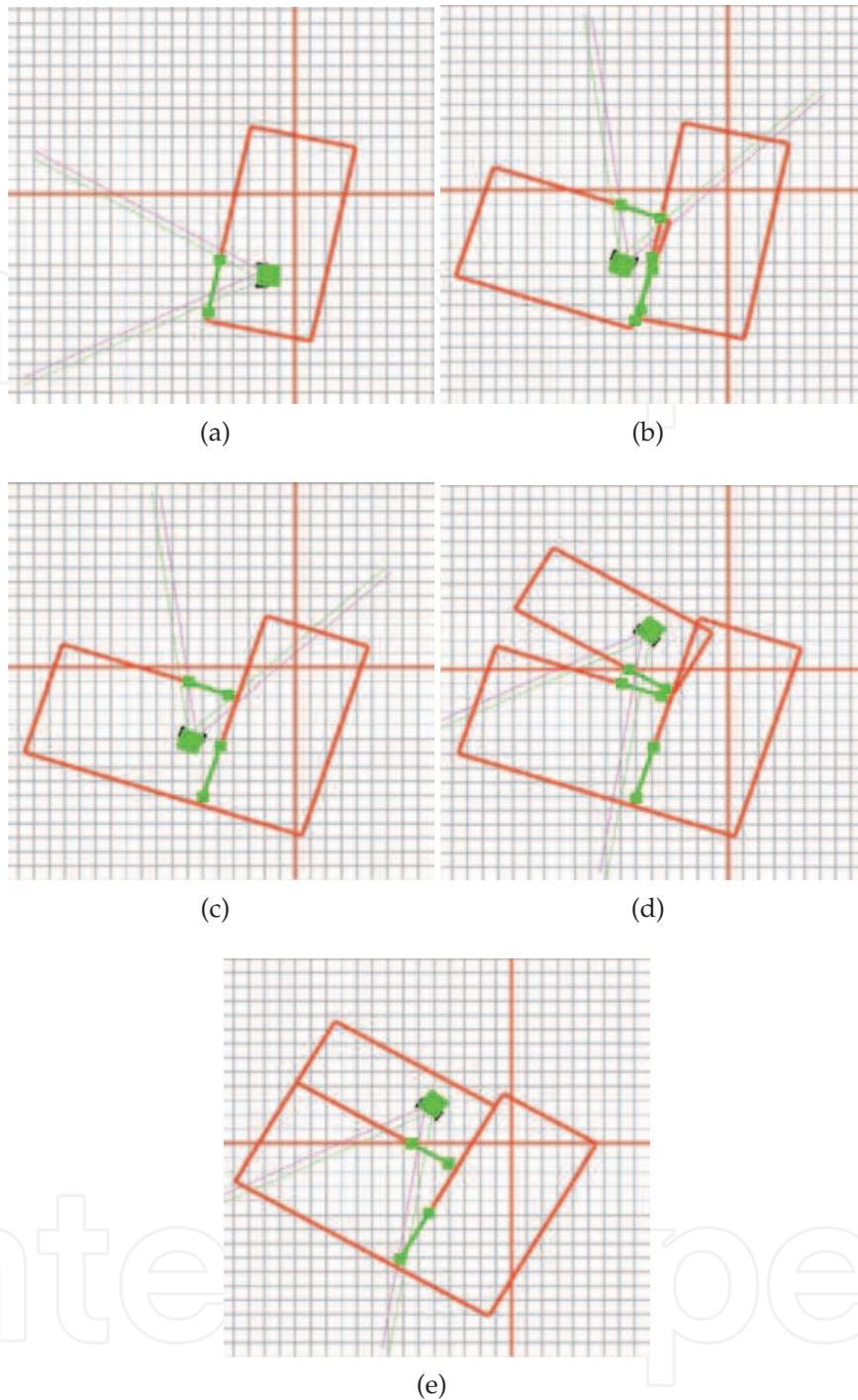
Fig. 17. Modeling process during an autonomous exploration of the scene of figure 18.

Work in order to relax the *rectangle assumption*, allowing the robot to work with more general models such as *polylines*, is currently in progress. We are also studying the advantages of using formal grammars for topological modeling. In addition, we are also improving and testing the system for much bigger and cluttered environments.
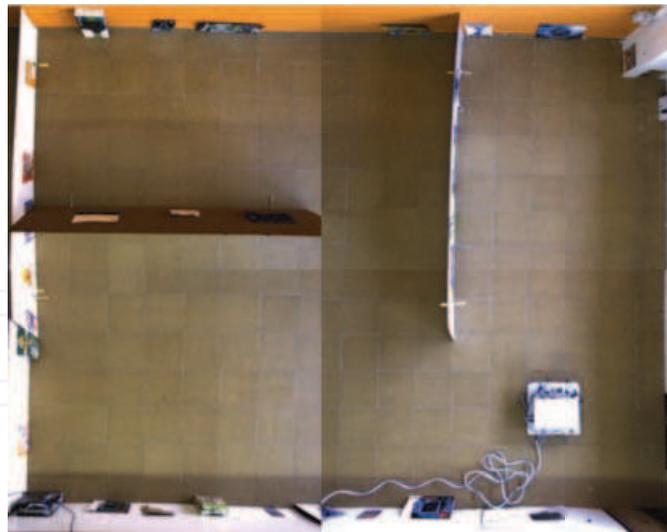
Fig. 18. Overhead view of the real scene of the experiment of figure 17.

7. Acknowledgements

8. References

Allport, A. (1987).   Selection for action: Some behavioral and neurophysiological considerations of attention and action, *in* H. Heuer & A. Sanders (eds), *Perspectives on perception and action*, Erlbaum.

Bachiller, P., Bustos, P. & Manso, L. (2008). Attentional selection for action in mobile robots, *in* J. Aramburo & A. R. Trevino (eds), *Advances in robotics, automation and control*, InTech, pp. 111–136.

Duda, R. & Hart, P. (1972). Use of the hough transformation to detect lines and curves in pictures, *Commun. ACM* 15: 11–15.

Enright, J. (1998). Monocularly programmed human saccades during vergence changes?, *Journal of Physiology* 512: 235–250.

Frintrop, S. (2006). *VOCUS: A Visual Attention System for Object Detection and Goal-Directed Search*, Vol. 3899 of *Lecture Notes in Computer Science*, Springer.

Itti, L. & Koch, C. (2000). A saliency-based search mechanism for overt and covert shifts of visual attention, *Vision Research* 40: 1489–1506.

Jung, C. & Schramm, R. (2004). Rectangle detection based on a windowed hough transform, *Proceedins of the XVII Brasilian Symposium on Computer Graphics and Image Processing*, pp. 113–120.

Koch, C. & Ullman, S. (1985). Shifts in selective visual attention: towards the underlying neural circuitry, *Human Neurobiology* 4: 219–227.

Lagunovsky, D. & Ablameyko, S. (1999). Straight-line-based primitive extraction in grey-scale object recognition, *Pattern Recognition Letters* 20(10): 1005–1014.

Lin, C. & Nevatia, R. (1998). Building detection and description from a single intensity image, *Computer Vision and Image Understanding* 72(2): 101–121.

Matas, J., Galambos, C. & Kittler, J. (2000). Robust detection of lines using the progressive probabilistic hough transform, *Computer Vision and Image Understanding* 78(1): 119–137.

Mateos, J., Sánchez-Domínguez, A., Manso, L., Bachiller, P. & Bustos, P. (2010). Robex: an open-hardware robotics platform, *Workshop of Physical Agents*.

Montijano, E. & Sagues, C. (2009). Topological maps based on graphs of planar regions, *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pp. 1661–1666.

Navalpakkam, V. & Itti, L. (2006). An integrated model of top-down and bottom-up attention for optimizing detection speed, *CVPR '06: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, pp. 2049–2056.

Olson, E. (2008). *Robust and Efficient Robotic Mapping*, PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, USA.

Palmer, P., Kittler, J. & Petrou, M. (1994). Using focus of attention with the hough transform for accurate line parameter estimation, *Pattern Recognition* 27(9): 1127–1134.

Robbins, H. & Monro, S. (1951). A stochastic approximation method, *The Annals of Mathematical Statistics* 22(3): 400–407.

Rosenfeld, A. (1969). Picture processing by computer, *ACM Comput. Surv.* 1: 147–176.

Simhon, S. & Dudek, G. (1998). A global topological map formed by local metric maps, *In IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 1708–1714.

Tao, W.-B., Tian, J.-W. & Liu, J. (2002). A new approach to extract rectangular building from aerial urban images, *Signal Processing, 2002 6th International Conference on*, Vol. 1, pp. 143 – 146.

Thrun, S. (1998). Learning metric-topological maps for indoor mobile robot navigation, *Artificial Intelligence* 99(1): 21–71.

Tomatis, N., Nourbakhsh, I. & Siegwart, R. (2003). Hybrid simultaneous localization and map building: a natural integration of topological and metric, *Robotics and Autonomous Systems* 44(1): 3–14.

Torralba, A., Oliva, A., Castelhano, M. S. & Henderson, J. M. (2006). Contextual guidance of eye movements and attention in real-world scenes: the role of global features in object search., *Psychol Rev* 113: 766–786.

Van Zwynsvoorde, D., Simeon, T. & Alami, R. (2000). Incremental topological modeling using local voronoï-like graphs, *Proc. of IEEE/RSJ Int. Conf. on Intelligent Robots and System (IROS 2000)*, Vol. 2, pp. 897–902.

Yan, F., Zhuang, Y. & Wang, W. (2006). Large-scale topological environmental model based particle filters for mobile robot indoor localization, *Robotics and Biomimetics, IEEE International Conference on* 0: 858–863.

Zhu, Y., Carragher, B., Mouche, F. & Potter, C. (2003). Automatic particle detection through efficient hough transforms, *IEEE Trans. Med. Imaging* 22(9).

Advances in Stereo Vision

Edited by Prof. Jose R.A. Torreao

Stereopsis is a vision process whose geometrical foundation has been known for a long time, ever since the experiments by Wheatstone, in the 19th century. Nevertheless, its inner workings in biological organisms, as well as its emulation by computer systems, have proven elusive, and stereo vision remains a very active and challenging area of research nowadays. In this volume we have attempted to present a limited but relevant sample of the work being carried out in stereo vision, covering significant aspects both from the applied and from the theoretical standpoints.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Pilar Bachiller, Pablo Bustos and Luis Manso (2011). Attentional Behaviors for Environment Modeling by a Mobile Robot, Advances in Stereo Vision, Prof. Jose R.A. Torreao (Ed.), ISBN: 978-953-307-837-3, InTech, Available from: http://www.intechopen.com/books/advances-in-stereo-vision/attentional-behaviors-for-environment-modeling-by-a-mobile-robot

INTECH
open science | open minds