

# Visual Based Localization for a Legged Robot<sup>\*</sup>

Francisco Martín, Vicente Matellán, Jose María Cañas, Pablo Barrera

Robotic Labs (GSyC), ESCET, Universidad Rey Juan Carlos,  
C/ Tulipán s/n CP. 28933 Móstoles (Madrid), Spain.  
{fmartin,vmo,jmplaza,barrera}@gsyc.escet.urjc.es

**Abstract.** This paper presents a visual based localization mechanism for a legged robot in indoor office environments. Our proposal is a probabilistic approach which uses partially observable Markov decision processes. We use a precompiled topological map where natural landmarks like doors or ceiling lights are recognized by the robot using its on-board camera. Experiments have been conducted using the AIBO Sony robotic dog showing that it is able to deal with noisy sensors like vision and to approximate world models representing indoor office environments. The major contributions of this work is the use of an *active vision* as the main input and localization in *not-engineered environments*.

## 1 Introduction

One of the basic tools for mobile robot operations is the localization capability [6]. It can be defined as the ability of a robot to determine its position in a map using its own sensors. Many works [13] have been developed to estimate the robot location as robot behavior may depend on its position inside the world.

Most of these algorithms, i.e. [11], have been designed for robots equipped with wheels, where locally accurate odometric information can be achieved and 360° sensory information is available. These requirement makes these methods unusable in legged robots. The solution we present is intended to solve the problem for a legged robot where odometric information is not reliable, even locally, and 360° sensory information is not available. This type of information from sonar or laser sensors is easier to process than camera images, available in our robot. The camera swinging in legged robots don't let a continuous image processing, making unusable the majority of the wheeled robot techniques for navigation, for instance.

In the literature, some works face this problem using vision as main sensor [3],[2] and [14], but most of them make their experiments in reduced and engineered environment (that is, placing ad-hoc landmarks or active beacons), mainly in the Robocup four legged league. In contrast, our work has been tested in a large office environment using natural pre-existing landmarks (doors, ceiling lights, etc). In addition, most of the approaches using vision as the main sensor

---

<sup>\*</sup> This work has been supported by grant DPI2004-07993-C03-01 of Spanish Government

for localization are *passive*, i.e. neither the sensor position, nor orientation are controlled. Our approach is *active*, as long as it commands the sensor orientation to get the information we need from the environment.

Another key difference with previous works is that our approximation is topological and most of the previously works use a metric approach. The office environment has been divided into states, where the set of nodes is built depending on the observations that can be obtained in each place of this environment (flat corridors and foyers). Others works using topological localization are, for instance [10], but again, these approaches are used in wheeled robots, where the information needed for applying these techniques are not available.

Our work is based on Partially Observable Markov Decision Processes [3]. We calculate a probability density (belief) over the entire states space (nodes of the topological map). These technique is also used in many other works as [1], [4] and [9], but using ultrasonic or infrared sensors to determine the obstacles around the robot, and using available odometry information. In [7], a markovian vision-based method is used, but the information extracted from the images are histograms and scale-invariant (SIFT) features [8] that are calculated from a wide set of image obtained in each map location. Our approach do not need previously taken images, only a 2D map where the landmarks needed for our model are displayed.

There are other approaches which use sampling algorithms for localization. In [12] *Monte Carlo* approximation is used, but in very controlled scenarios, and in [5] it has been reported that this technique is not effective in noisy environments as ours. Once again, these works have been tested in reduced and engineered environments (RoboCup mainly).

Summarizing, the major contributions of this paper are three:

1. The development of a probabilistic localization technique for legged robots. It has been tested in large environments. Other works previously cited using same robots are mainly devoted to small engineered spaces (mainly the Robocup playground).
2. The development of a topological framework for navigation in not-engineered indoor environments. Majority of works on legged robots localization has focused on metric localization in engineered environments.
3. The use of the robot on-board camera as the main sensor for existing landmarks (doors, ceiling lights, etc.) detection in natural indoor scenarios, and the control of the position of the camera.

Our work has been developed in an AIBO ERS7 robot. This robot is a completely autonomous robot which incorporates an embedded MIPS processor running at 576MHz, and 64MB of main memory. It gets information from the environment through a 350K-pixel color camera and 2 infrared sensors. AIBO locomotion main characteristic is its dog aspect with four legs.

The remainder of this paper is organized as follows: in section 2 we make a brief review of the Markov Localization technique used. In section 3 we describe our model and its components in detail. In section 4 the experiments and results

are shown. Finally, we expose some conclusions and envisioned improvements in section 5.

## 2 Markovian localization framework

Localization based on indirect information provided by the robot sensors (sonar, laser, etc.) has been successfully integrated in the probabilistic framework and has exhibited good results [13]. In particular, sampling methods that speed up the estimation [3] are currently the most popular ones [11].

In our work, we have used a partially observable Markov decision processes (POMDP) where a probability distribution  $Bel$ , over all the possible locations  $S = \{s_1, s_2, \dots\}$  is maintained.  $Bel_t(S = s)$  represents the belief of being in state  $s$  at time  $t$ . Depending on the knowledge about the initial position of the robot  $Bel_0(S)$ , the initial state will be uniformly distributed, if the position is not known. If the position is known, the distribution will be centered in the initial state.

The belief  $Bel$  actualization is divided in two atomic steps:

**Movement step.** Robot actions are modelled by the probability  $p(s'|s, a)$  (action model). This is the probability of reaching state  $s'$  if an action  $a$  is executed at state  $s$ . To obtain the a priori belief for the whole set of states  $Bel_t(S')$  bayesian updating is assumed. When an action is executed we apply:

$$Bel_t(s') = \sum_{s \in S} p(s'|s, a) \cdot Bel_{t-1}(s), \forall s' \in S \quad (1)$$

**Observation step.** To calculate the corrected belief  $Bel_t(S)$  we take  $p(o|s)$  (sensor model) as the probability of getting the observation  $o$  being in the state  $s$  and we operate, as it is described in [9]. When a new independent observation is obtained the belief is updated using (2).

$$Bel_t(s) = p(o|s) \cdot Bel_t(s'), \forall s, s' \in S \quad (2)$$

## 3 Our Model

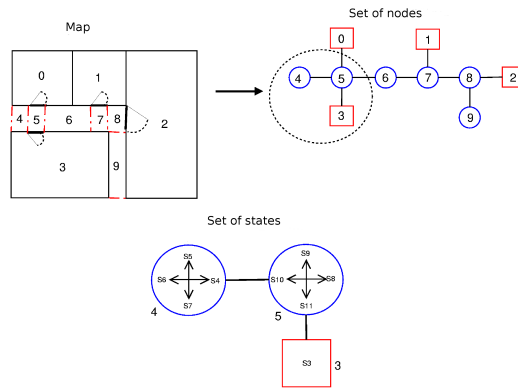
Summarizing, our localization method needs three components to be described, and these will be defined in more detail in this section:

1. The map and how it is translated to a set of states.
2. A set of actions that the robot can perform and their probabilistic action model related to states  $p(s'|s, a)$
3. A set of observations the the robot perceives from the environment, and its probabilistic model related to states  $p(o|s)$

### 3.1 The state space

We name possible locations of the robot as “states”. These states are defined over an indoor office environment (see Fig. 1) made up by corridors (represented in the set of nodes as circles) and rooms (represented as boxes). Nodes are defined as places with similar sensory characteristics.

Once the set of nodes has been defined, each node has been divided in four different states, representing the same robot position with four orientations: north, east, south, and west.



**Fig. 1.** The set of states is built from the map and its node decomposition

### 3.2 Action model

The action primitives we have implemented in this work are: to turn  $90^\circ$  on the left  $a_{\{TL\}}$ , to turn  $90^\circ$  on the right  $a_{\{TR\}}$  and go forward  $a_{\{F\}}$  until the next state with the same orientation is reached.

When the robot executes an action primitive, i.e. when the robot moves, it updates the belief as it is shown in (3). The action model defines  $p(s'|s, a)$  as the probability of to reach state  $s'$ , starting at state  $s$  and executing the action  $a$ :

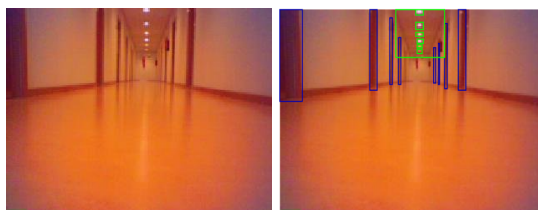
$$p(s'|s, a), \forall s \in S, \forall a \in A = \{a_{\{F\}}, a_{\{TL\}}, a_{\{TR\}}\} \quad (3)$$

This probability  $p(s'|s, a)$  will represent our action model and it is calculated *a priori*, depending on the possible action the robot can perform in that state space.

### 3.3 Sensor Model

Our sensor model take three types of sensations from the image taken by the robot’s camera:

**Depth.** The main target for this observation is measure how far the robot is from the wall when it is orientated to the end of the corridor. For this purpose we detect the number of ceiling lights that the robot perceive. If the number of ceiling lights is high, the robot is far from the end. If this measure is low, the robot is near to the end. In Fig. 3.3 we can see the original image and the image with the ceiling lights and doors detected.



**Fig. 2.** Image information extraction results. Detecting 6 ceiling lights and 8 doors

**Doors.** Using a color filter, the robot is able to count the number of doors it can observe ahead.

**Near landmark.** This observation give us information about which landmarks are around the robot. We define landmarks as the doors or walls that are situated in the right, left and front side of robot.

Once the data is collected, we apply the equation (2) to correct the belief.

$$Bel_{subsequent}(s) = p(\mathbf{o}|s) \cdot Bel_{previous}(s), \forall s \in S \quad (4)$$

## 4 Experiments

We have made several experiments in a corridor of our office environment in a normal daily work. In Fig. 3.3, we can see the corridor that we have used for the experiment and how we have divided it into nodes. Afterwards we have divided the set of nodes into states. This office environment is very symmetric and that is why this scenario entrails much more difficulty for the localization system.

For the experimental results, we have used the error function shown in equation (5), where  $state_{high}$  denotes the state with the greatest belief and  $state_{current}$  is the robot actual position. The  $distance$  is measured as the number of steps needed to reach one state from another.

$$error = ||prob(state_{high} - prob(state_{current})) \cdot distance(state_{high}, state_{current})|| \quad (5)$$

#### 4.1 Ability to recover of an action error

In the first experiment we want to verify if the system is robust enough to cope with action errors. The system must be able to detect when the movement was wrong using its sensors, and recover from this situation.

For this purpose, we have situated the robot in state 15 and we ordered it to go forward along the corridor. The robot knows its initial position, in other words, the probability distribution is concentrated in the state 15.

$$Bel_0(s_{15}) = 1$$

$$Bel_0(s_i) = 0, \forall s_i \in S, s_i \neq s_{15}$$

We commanded the robot to perform the actions sequentially. Some of these actions did not execute correctly. Although in the *movement step* the belief is changed in a wrong way, in the *observation step* the belief is corrected due to the information obtained from the environment in all the experiments.

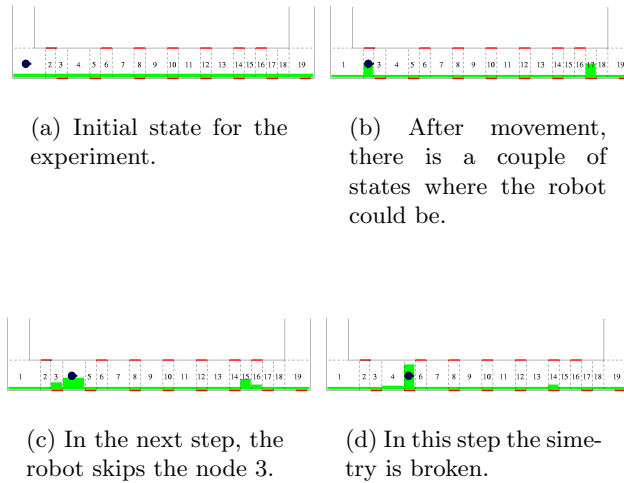
#### 4.2 Localization speed

In this experiment the robot does not know its starting position, so the first time the location probability distribution is uniform.

$$Bel_0(s_i) = \frac{1}{|S|}, \forall s_i \in S \quad (6)$$

This experiment was realized with a lot of sensor noise because there were a lot of people walking along the corridor. Despite this difficulty, the robot is able to be localized with a small error in a few movements and can recovery to sensor error quickly, as we see in Fig. 3(a)-3(d).

In Fig. 3(a) the robot starts at node one and the distribution (painted in green) is uniform along all the nodes. For this explanation we will talk about *node* instead of *states*, which is actually what we use in our model, to simplify this explanation. So, a node will be padded in green depending on the belief of the state situated in this node, orientated on the right. When the robot moves forward it reach to the node 2 (Fig. 3(b)) and it takes data from its sensors. With this data the model evolves and the probability is concentrated in state 2 and 17, because these two states have almost the same observation properties. The robot goes forward, but an error occurs and the robot reaches node 4, instead of node 3. This anomaly is observed in the model and it is corrected in the *observation phase*, as we see in Fig. 3(c). In the last movement the robot reaches the node 5 and then the simetry is broken, concentrating the probability in the node 5, as we see in Fig. 3(d).



**Fig. 3.** Experiment done in a corridor. The amount of green in each state represent the belief in it.

## 5 Conclusions

In this article we have presented the preliminary results for the localization of legged AIBO robots in not-engineered environments, using the vision as an active input sensor. We have shown that the robot is able to localize in real time itself even in environments with noise produced by the human activity in a real office. It deal with uncertainly in its action and uses perceived natural landmarks of the environment as the main sensor input.

The data obtained from sensors, mainly the camera, is discriminant enough and let a fast convergence from an initial unknown state, where the belief over the set of states has been distributed uniformly. Also we have shown that the robot can overcome action failures while localizing, and it recovers from them in a efficient way.

The set of observations we have chosen have been descriptive enough to be efficient in the localization process. We think that the way we determine the number of doors and ceiling lights has been the key for the success of the localization system.

We believe that probabilistic navigation techniques hold great promise for getting legged robots reliable enough to operate in real office environments.

## References

- [1] Anthony R. Cassandra, Leslie Pack Kaelbling, and James A. Kurien. Acting under uncertainty: Discrete bayesian models for mobile robot navigation. In *Proceedings of IEEE/RSJ International Conference on Intelligent Robots and Systems*, 1996.
- [2] Stefan Enderle, Marcus Ritter, Dieter Fox, Stefan Sablatnög, Gerhard Kraetschmar, and Günther Palm. Soccer robot localization sporadic visual features. In *IAS-00*, 2000.
- [3] Dieter Fox, Wolfram Burgard, and Sebastian Thrun. Markov localization for mobile robots in dynamic environments. *Journal of Artificial Intelligence Research*, 1999.
- [4] Franck Gechter, Thomas Vincent, and François Charpillet. Robot localization by stochastic vision based device. In *Cybernetics and Informatics (SCI)*, 2001.
- [5] Pablo Guerrero and Javier Ruiz del Solar. Auto-localización de un robot móvil aibo mediante el método de monte carlo. Technical report, Instituto de Ingenieros de Chile, 2003.
- [6] J.Borenstein, B.Everett, and L.Feng. *Navigating mobile robots: Systems and techniques*. Ltd. Wesley, MA, 1996.
- [7] Jana Kosecká and Fayin li. Vision based topological markov localization. In *Proceedings of the 2004 IEEE International Conference on Robotics and Automation*, Barcelona (Spain), April 2004.
- [8] David G. Lowe. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision*, 60(2):91–110, 2004.
- [9] María E. López, Luis Miguel Bergasa, and M.S.Escudero. Visually augmented pomdp for indoor robot navigation. In *Applied Informatics*, pages 183–187, 2003.
- [10] Dandapani Radhakrishnan and Illah Nourbakhsh. Topological localization by training a vision-based transition detector. *IROS*, 1999.
- [11] Reid Simmons and Sven Koenig. Probabilistic navigation in partially observable environments. In *Proceedings of the 1995 International Joint Conference on Artificial Intelligence*, Montreal (Canada), July 1995.
- [12] Mohan Sridharan, Gregory Kuhlmann, and Peter Stone. Practical vision-based monte carlo localization on a legged robot. In *IEEE International Conference on Robotics and Automation*, April 2005.
- [13] S.Thrun. Robotic mapping: A survey. *Technical Report CMU-CS-02-111*, 2002.
- [14] Manuela Veloso, Elly Winner, Scott Lenser, James Bruce, and Tucker Balch. Vision-servoed localization and behavior-based planning for an autonomous quadruped legged robot. In *Proceedings of the Fifth International Conference on Artificial Intelligence Planning Systems*, pages 387–394, Breckenridge, CO, April 2000.